

9/23/02

09/677,822

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) Publication number:

0 251 446 B1

(12)

EUROPEAN PATENT SPECIFICATION

- (45) Date of publication of patent specification: **28.12.94** (51) Int. Cl.⁵: **C12N 15/00, C12N 9/54, C12N 1/00**
- (21) Application number: **87303761.8**
- (22) Date of filing: **28.04.87**

- (54) **Non-human Carbonyl hydrolase mutants, DNA sequences and vectors encoding same and hosts transformed with said vectors.**

- (30) Priority: **30.04.86 US 858594**
06.04.87 US 35652
- (43) Date of publication of application:
07.01.88 Bulletin 88/01
- (45) Publication of the grant of the patent:
28.12.94 Bulletin 94/52
- (64) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE
- (56) References cited:
EP-A- 0 130 756
WO-A-87/04461
WO-A-87/05050

**ABSTRACTS OF THE 190TH AMERICAN
CHEMICAL SOCIETY NATIONAL MEETING,
vol. 190,1985, page 23, no. 47; R.R. BOTT et
al.: "Protein engineering of subtilisin"**

- (73) Proprietor: **GENENCOR INTERNATIONAL, INC.**
180 Kimball Way
South San Francisco, CA 94080 (US)
- (72) Inventor: **Wells, James Allen**
64 Otay Avenue
San Mateo
CA 94403 (US)
Inventor: **Cunningham, Brian C.**
24 Olive Avenue
Piedmont
CA 94611 (US)
Inventor: **Caldwell, Robert Mark**
1828 Broadway
No.101
San Francisco
Ca 94109 (US)
Inventor: **Bott, Richard Ray**
3032 Hillside drive
Burlingame
CA 94010 (US)

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

EP 0 251 446 B1

JOURNAL OF CELLULAR BIOCHEMISTRY SUPPL., vol. 0, no. 10, part A, 1986, page 271, no. E101, SYMPOSIUM ON PROTEASES IN BIOLOGICAL CONTROL AND BIOTECHNOLOGY, 15th ANNUAL UCLA, MEETING ON MOLECULAR AND CELLULAR BIOLOGY, Los Angeles, CA., 9th-15th February 1986; P. BRYAN et al.: "Protein engineering of subtilisin proteases of enhanced stability"

WORLD BIOTECH. REPORT, vol. 2, 1985, pages 51-59, Online Publications, Pinner, GB; R. BOTT: "Modeling & crystallographic analysis of site-specific mutants of subtilisin"

JOURNAL OF CELLULAR BIOCHEMISTRY SUPPL., vol. 0, no. 11, part C, 1987, page 200, no. N024, New York, US; D.A. ESTELL et al.: "Tailoring enzymatic properties through multiple mutations"

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 84, March 1987, pages 1219-1223, Washington, D.C., US; J.A. WELLS et al.: "Designing substrate specificity by protein engineering of electrostatic interactions"

BIOCHEMISTRY, vol. 26, no. 8, April 1987, pages 2077-2082, American Chemical Society, Washington, D.C., US; M.W. PANTOLIANO et al.: "Protein engineering of subtilisin BPN': enhanced stabilization through the introduction of two cysteines to form a disulfide bond"

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 83, June 1986, pages 3743-3745, Washington, D.C., US; P. BRYAN et al.: "Site-directed mutagenesis and the role of the oxyanion hole in subtilisin"

NATURE, vol. 318, 28th November 1985, pages 375-376, London, GB; P.G. THOMAS et al.: "Tailoring the pH dependence of enzyme catalysis using protein engineering"

JOURNAL OF BACTERIOLOGY, vol. 158, no. 2, May 1984, pages 411-418, American Society for Microbiology, Washington, D.C., US; M.L. STAHL et al.: "Replacement of the Bacillus subtilis subtilisin structural gene with an in vitro-derived deletion mutation"

Inventor: Estell, David Aaron
250 Diablo Avenue
Mountain View
CA 94043 (US)
Inventor: Power, Scott Douglas
732 Olive Court
San Bruno
CA 94066 (US)

(74) Representative: Armitage, Ian Michael et al
MEWBURN ELLIS
York House
23 Kingsway
London WC2B 6HP (GB)

NUCLEIC ACIDS RESEARCH, vol. 11, no. 22,
November 1983, pages 7911-7925, IRL Press
Ltd, Cambridge, GB; J.A. WELLS et al.: "Clon-
ing, sequencing, and secretion of Bacillus
amyloliquefaciens subtilisin in Bacillus sub-
tilis"

Description

The recent development of various *in vitro* techniques to manipulate the DNA sequences encoding naturally-occurring polypeptides as well as recent developments in the chemical synthesis of relatively short sequences of single and double stranded DNA has resulted in the speculation that such techniques can be used to modify enzymes to improve some functional property in a predictable way. Ulmer, K.M. (1983) *Science* 219, 666-671. The only working example disclosed therein is the substitution of a single amino acid within the active site of tyrosyl-tRNA synthetase (Cys35→Ser) which lead to a reduction in enzymatic activity. See Winter, G., et al. (1982) *Nature* 299, 756-758; and Wilkinson, A.J., et al. (1983) *Biochemistry* 22, 3581-3586 (Cys35→Gly mutation also resulted in decreased activity).

When the same t-RNA synthetase was modified by substituting a different amino acid residue within the active site with two different amino acids, one of the mutants (Thr51→Ala) reportedly demonstrated a predicted moderate increase in *k_{cat}/K_m* whereas a second mutant (Thr51→Pro) demonstrated a massive increase in *k_{cat}/K_m* which could not be explained with certainty. Wilkinson, A.H., et al. (1984) *Nature* 307, 187-188.

Another reported example of a single substitution of an amino acid residue is the substitution of cysteine for isoleucine at the third residue of T4 lysozyme. Perry, L.J., et al. (1984) *Science* 226, 555-557. The resultant mutant lysozyme was mildly oxidized to form a disulfide bond between the new cysteine residue at position 3 and the native cysteine at position 97. This crosslinked mutant was initially described by the author as being enzymatically identical to, but more thermally stable than, the wild type enzyme. However, in a "Note Added in Proof", the author indicated that the enhanced stability observed was probably due to a chemical modification of cysteine at residue 54 since the mutant lysozyme with a free thiol at Cys54 has a thermal stability identical to the wild type lysozyme.

Similarly, a modified dihydrofolate reductase from *E.coli* has been reported to be modified by similar methods to introduce a cysteine which could be cross linked with a naturally-occurring cysteine in the reductase. Villafranca, D.E., et al. (1983) *Science* 222, 782-788. The author indicates that this mutant is fully reactive in the reduced state but has significantly diminished activity in the oxidized state. In addition, two other substitutions of specific amino acid residues are reported which resulted in mutants which had diminished or no activity.

EPO Publication No. 0130756 discloses the substitution of specific residues within *B. amyloliquefaciens* subtilisin with specific amino acids. Thus, Met222 has been substituted with all 19 other amino acids, Gly166 with 9 different amino acids and Gly169 with Ala and Ser.

As set forth below, several laboratories have also reported the use of site directed mutagenesis to produce the mutation of more than one amino acid residue within a polypeptide.

The amino-terminal region of the signal peptide of the prolipoprotein of the *E. coli* outer membrane was stated to be altered by the substitution or deletion of residues 2 and 3 to produce a charge change in that region of the polypeptide. Inouye, S., et al. (1982) *Proc. Nat. Acad. Sci. USA* 79, 3438-3441. The same laboratory also reported the substitution and deletion of amino acid residues 9 and 14 to determine the effects of such substitution on the hydrophobic region of the same signal sequence. Inouye, S., et al. (1984) *J. Biol. Chem.* 259, 3729-3733.

Double mutants in the active site of tyrosyl-t-RNA synthetase have also been reported. Carter, P.J., et al. (1984) *Cell* 38, 835-840. In this report, the improved affinity of the previously described Thr51→Pro mutant for ATP was probed by producing a second mutation in the active site of the enzyme. One of the double mutants, Gly35/Pro51, reportedly demonstrated an unexpected result in that it bound ATP in the transition state better than was expected from the two single mutants. Moreover, the author warns, at least for one double mutant, that it is not readily predictable how one substitution alters the effect caused by the other substitution and that care must be taken in interpreting such substitutions.

A mutant is disclosed in U.S. Patent No. 4,532,207, wherein a polyarginine tail was attached to the C-terminal residue of β -urogastrone by modifying the DNA sequence encoding the polypeptide. As disclosed, the polyarginine tail changed the electrophoretic mobility of the urogastrone-polyarginine hybrid permitting selective purification. The polyarginine was subsequently removed, according to the patentee, by a polyarginine specific exopeptidase to produce the purified urogastrone. Properly construed, this reference discloses hybrid polypeptides which do not constitute mutant polypeptides containing the substitution, insertion or deletion of one or more amino acids of a naturally occurring polypeptide.

Single and double mutants of rat pancreatic trypsin have also been reported. Craik, C.S., et al. (1985) *Science* 228, 291-297. As reported, glycine residues at positions 216 and 226 were replaced with alanine residues to produce three trypsin mutants (two single mutants and one double mutant). In the case of the single mutants, the authors stated expectation was to observe a differential effect on *K_m*. They instead

reported a change in specificity (kcat/Km) which was primarily the result of a decrease in kcat. In contrast, the double mutant reportedly demonstrated a differential increase in Km for lysyl and arginyl substrates as compared to wild type trypsin but had virtually no catalytic activity.

The references discussed above are provided solely for their disclosure prior to the filing date of the instant case, and nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or priority based on earlier filed applications.

Based on the above references, however, it is apparent that the modification of the amino acid sequence of wild type enzymes often results in the decrease or destruction of biological activity.

Accordingly, it is an object herein to provide carbonyl hydrolase mutants which have at least one property which is different from the same property of the carbonyl hydrolase precursor from which the amino acid of said mutant is derived.

It is a further object to provide mutant DNA sequences encoding such carbonyl hydrolase mutants as well as expression vectors containing such mutant DNA sequences.

Still further, another object of the present invention is to provide host cells transformed with such vectors as well as host cells which are capable of expressing such mutants either intracellularly or extracellularly.

Summary of the Invention

The invention includes carbonyl hydrolase mutants, preferably having at least one property which is substantially different from the same property of the precursor non-human carbonyl hydrolase from which the amino acid sequence of the mutant is derived. These properties include oxidative stability, substrate, specificity catalytic activity, thermal stability, alkaline stability, pH activity profile and resistance to proteolytic degradation. The precursor carbonyl hydrolase may be naturally occurring carbonyl hydrolases or recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is derived by the substitution, deletion or insertion of one or more amino acids of the precursor carbonyl hydrolase amino acid sequence.

The invention also includes mutant DNA sequences encoding such carbonyl hydrolase mutants. Further the invention includes expression vectors containing such mutant DNA sequences as well as host cells transformed with such vectors which are capable of expressing said carbonyl hydrolase mutants.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence of the coding strand, correlated with the amino acid sequence of *B. amyloliquefaciens* subtilisin gene. Promoter (p) ribosome binding site (rbs) and termination (term) regions of the DNA sequence as well as sequences encoding the presequence (PRE) putative prosequence (PRO) and mature form (MAT) of the hydrolase are also shown.

Figure 2 is a schematic diagram showing the substrate binding cleft of subtilisin together with substrate.

Figure 3 is a stereo view of the S-1 binding subsite of *B. amyloliquefaciens* subtilisin showing a lysine P-1 substrate bound in the site in two different ways. Figure 3A shows Lysine P-1 substrate bound to form a salt bridge with a Glu at position 156. Figure 3B shows Lysine P-1 substrate bound to form a salt bridge with Glu at position 166.

Figure 4 is a schematic diagram of the active site of subtilisin Asp32, His64 and Ser221.

Figures 5A and 5B depict the amino acid sequence of subtilisin obtained from various sources. The residues directly beneath each residue of *B. amyloliquefaciens* subtilisin are equivalent residues which (1) can be mutated in a similar manner to that described for *B. amyloliquefaciens* subtilisin, or (2) can be used as a replacement amino acid residue in *B. amyloliquefaciens* subtilisin. Figure 5C depicts conserved residues of *B. amyloliquefaciens* subtilisin when compared to other subtilisin sequences.

Figures 6A and 6B depict the inactivation of the mutants Met222L and Met222Q when exposed to various organic oxidants.

Figure 7 depicts the ultraviolet spectrum of Met222F subtilisin and the difference spectrum generated after inactivation by diperdodecanoic acid (DPDA).

Figure 8 shows the pattern of cyanogen bromide digests of untreated and DPDA oxidized subtilisin Met222F on high resolution SDS-pyridine peptide gels.

Figure 9 depicts a map of the cyanogen bromide fragments of Fig. 8 and their alignment with the sequence of subtilisin Met222F.

Figure 10 depicts the construction of mutations between codons 45 and 50 of *B. amyloliquefaciens* subtilisin.

Figure 11 depicts the construction of mutations between codons 122 and 127 of B. amyloliquefaciens subtilisin.

Figure 12 depicts the effect of DPDA on the activity of subtilisin mutants at positions 50 and 124 in subtilisin Met222F.

5 Figure 13 depicts the construction of mutations at codon 166 of B. amyloliquefaciens subtilisin.

Figure 14 depicts the effect of hydrophobicity of the P-1 substrate side-chain on the kinetic parameters of wild-type B. amyloliquefaciens subtilisin.

Figure 15 depicts the effect of position 166 side-chain substitutions on P-1 substrate specificity. Figure 15A shows position 166 mutant subtilisins containing non-branched alkyl and aromatic side-chain substitutions arranged in order of increasing molecular volume. Figure 15B shows a series of mutant enzymes
10 progressing through β - and γ -branched aliphatic side chain substitutions of increasing molecular volume.

Figure 16 depicts the effect of position 166 side-chain volume on log kcat/Km for various P-1 substrates.

Figure 17 shows the substrate specificity differences between Ile166 and wild-type (Gly166) B. amyloliquefaciens subtilisin against a series of aliphatic and aromatic substrates. Each bar represents the
15 difference in log kcat/Km for Ile166 minus wild-type (Gly166) subtilisin.

Figure 18 depicts the construction of mutations at codon 169 of B. amyloliquefaciens subtilisin.

Figure 19 depicts the construction of mutations at codon 104 of B. amyloliquefaciens subtilisin.

Figure 20 depicts the construction of mutations at codon 152 B. amyloliquefaciens subtilisin.

20 Figure 21 depicts the construction of single mutations at codon 156 and double mutations at codons 156 and 166 of B. amyloliquefaciens subtilisin.

Figure 22 depicts the construction of mutations at codon 217 for B. amyloliquefaciens subtilisin.

Figure 23 depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

25 Figure 23A depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 24 depicts the kcat/Km versus pH profile for mutations at codon 222 in B. amyloliquefaciens subtilisin.

Figure 25 depicts the constructing mutants at codons 94, 95 and 96.

30 Figures 26 and 27 depict substrate specificity of various wild type and mutant subtilisins for different substrates.

Figures 28 A, B, C and D depict the effect of charge in the P-1 binding sites due to substitutions at codon 156 and 166.

Figures 29 A and B are a stereoview of the P-1 binding site of subtilisin BPN' showing a lysine P-1 substrate bound in the site in two ways. In 29A, Lysine P-1 substrate is built to form a salt bridge with a Glu at codon 156. In 29B, Lysine P-1 substrate is built to form a salt bridge with Glu at codon 166.
35

Figure 30 demonstrates residual enzyme activity versus temperature curves for purified wild-type (Panel A), C22/C87 (Panel B) and C24/C87 (Panel C).

Figure 31 depicts the strategy for producing point mutations in the subtilisin coding sequence by misincorporation of α -thioldeoxynucleotide triphosphates.
40

Figure 32 depicts the autolytic stability of purified wild type and mutant subtilisins 170E, 107V, 213R and 107V/213R at alkaline pH.

Figure 33 depicts the autolytic stability of purified wild type and mutant subtilisins V50, F50 and F50/V107/R213 at alkaline pH.

45 Figure 34 depicts the strategy for constructing plasmids containing random cassette mutagenesis over residues 197 through 228.

Figure 35 depicts the oligodeoxynucleotides used for random cassette mutagenesis over residues 197 through 228.

Figure 36 depicts the construction of mutants at codon 204.

50 Figure 37 depicts the oligodeoxynucleotides used for synthesizing mutants at codon 204.

Detailed Description

The inventors have discovered that various single and multiple in vitro mutations involving the
55 substitution, deletion or insertion of one or more amino acids within a non-human carbonyl hydrolase amino acid sequence can confer advantageous properties to such mutants when compared to the non-mutated carbonyl hydrolase.

Specifically, *B. amyloliquefaciens* subtilisin, an alkaline bacterial protease, has been mutated by modifying the DNA encoding the subtilisin to encode the substitution of one or more amino acids at various amino acid residues within the mature form of the subtilisin molecule. These *in vitro* mutant subtilisins have at least one property which is different when compared to the same property of the precursor subtilisin.

5 These modified properties fall into several categories including: oxidative stability, substrate specificity, thermal stability, alkaline stability, catalytic activity, pH activity profile, resistance to proteolytic degradation, Km, kcat and Km/kcat ratio.

Carbonyl hydrolases are enzymes which hydrolyze compounds containing

10



15

bonds in which X is oxygen or nitrogen. They include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. Naturally occurring carbonyl hydrolases principally include hydrolases, e.g. lipases and peptide hydrolases, e.g. subtilisins or metalloproteases. Peptide hydrolases include α -aminoacylpeptide hydrolase, peptidylamino-acid hydrolase, acylamino hydrolase, serine carboxypeptidase, 20 metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exoproteases.

"Recombinant carbonyl hydrolase" refers to a carbonyl hydrolase in which the DNA sequence encoding the naturally occurring carbonyl hydrolase is modified to produce a mutant DNA sequence which encodes the substitution, insertion or deletion of one or more amino acids in the carbonyl hydrolase amino acid sequence. Suitable modification methods are disclosed herein and in EPO Publication No. 0130756 25 published January 9, 1985.

Subtilisins are bacterial carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally occurring subtilisin or a recombinant subtilisin. A series of naturally occurring subtilisins is known to be produced and often secreted by various bacterial 30 species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the amino to carboxy terminus is aspartate-histidineserine. 35 In the chymotrypsin related proteases the relative order, however is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases.

"Recombinant subtilisin" refers to a subtilisin in which the DNA sequence encoding the subtilisin is modified to produce a mutant DNA sequence which encodes the substitution, deletion or insertion of one or 40 more amino acids in the naturally occurring subtilisin amino acid sequence. Suitable methods to produce such modification include those disclosed herein and in EPO Publication No. 0130756. For example, the subtilisin multiple mutant herein containing the substitution of methionine at amino acid residues 50, 124 and 222 with phenylalanine, isoleucine and glutamine, respectively, can be considered to be derived from the recombinant subtilisin containing the substitution of glutamine at residue 222 (Q222) disclosed in EPO 45 Publication No. 0130756. The multiple mutant thus is produced by the substitution of phenylalanine for methionine at residue 50 and isoleucine for methionine at residue 124 in the Q222 recombinant subtilisin.

"Carbonyl hydrolases" and their genes may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as *E. coli* or pseudomonas and gram positive bacteria such as micrococcus or bacillus. Examples of eucaryotic 50 organisms from which carbonyl hydrolase and their genes may be obtained include yeast such as *S. cerevisiae*, fungi such as *Aspergillus* sp., and non-human mammalian sources such as, for example, Bovine sp. from which the gene encoding the carbonyl hydrolase chymosin can be obtained. As with subtilisins, a series of carbonyl hydrolases can be obtained from various related species which have amino acid sequences which are not entirely homologous between the members of that series but which nevertheless 55 exhibit the same or similar type of biological activity. Thus, non-human carbonyl hydrolase as used herein has a functional definition which refers to carbonyl hydrolases which are associated, directly or indirectly, with procaryotic and non-human eucaryotic sources.

A "carbonyl hydrolase mutant" has an amino acid sequence which is derived from the amino acid sequence of a non-human "precursor carbonyl hydrolase". The precursor carbonyl hydrolases include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is "derived" from the precursor hydrolase amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the amino acid sequence of the precursor carbonyl hydrolase rather than manipulation of the precursor carbonyl hydrolase *per se*. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein and in EPO Publication No. 0130756.

Specific residues of *B. amyloliquefaciens* subtilisin are identified for substitution, insertion or deletion. These amino acid position numbers refer to those assigned to the *B. amyloliquefaciens* subtilisin sequence presented in Fig. 1. The invention, however, is not limited to the mutation of this particular subtilisin but extends to precursor carbonyl hydrolases containing amino acid residues which are "equivalent" to the particular identified residues in *B. amyloliquefaciens* subtilisin.

A residue (amino acid) of a precursor carbonyl hydrolase is equivalent to a residue of *B. amyloliquefaciens* subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analogous to a specific residue or portion of that residue in *B. amyloliquefaciens* subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor carbonyl hydrolase is directly compared to the *B. amyloliquefaciens* subtilisin primary sequence and particularly to a set of residues known to be invariant in all subtilisins for which sequence is known (Figure 5C). After aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of *B. amyloliquefaciens* subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 50% of conserved residues is also adequate to define equivalent residues. Conservation of the catalytic triad, Asp32/His64/Ser221 should be maintained.

For example, in Figure 5A the amino acid sequence of subtilisin from *B. amyloliquefaciens* *B. subtilisin* var. I168 and *B. licheniformis* (carlsbergensis) are aligned to provide the maximum amount of homology between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. These residues are identified in Fig. 5C.

These conserved residues thus may be used to define the corresponding equivalent amino acid residues of *B. amyloliquefaciens* subtilisin in other carbonyl hydrolases such as thermitase derived from *Thermoactinomyces*. These two particular sequences are aligned in Fig. 5B to produce the maximum homology of conserved residues. As can be seen there are a number of insertions and deletions in the thermitase sequence as compared to *B. amyloliquefaciens* subtilisin. Thus, in thermitase the equivalent amino acid of Tyr217 in *B. amyloliquefaciens* subtilisin is the particular lysine shown beneath Tyr217.

In Fig. 5A, the equivalent amino acid at position 217 in *B. amyloliquefaciens* subtilisin is Tyr. Likewise, in *B. subtilis* subtilisin position 217 is also occupied by Tyr but in *B. licheniformis* position 217 is occupied by Leu.

Thus, these particular residues in thermitase, and subtilisin from *B. subtilisin* and *B. licheniformis* may be substituted by a different amino acid to produce a mutant carbonyl hydrolase since they are equivalent in primary structure to Tyr217 in *B. amyloliquefaciens* subtilisin. Equivalent amino acids of course are not limited to those for Tyr217 but extend to any residue which is equivalent to a residue in *B. amyloliquefaciens* whether such residues are conserved or not.

Equivalent residues homologous at the level of tertiary structure for a precursor carbonyl hydrolase whose tertiary structure has been determined by x-ray crystallography, are defined as those for which the atomic coordinates of 2 or more of the main chain atoms of a particular amino acid residue of the precursor carbonyl hydrolase and *B. amyloliquefaciens* subtilisin (N on N, CA on CA, C on C, and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the carbonyl hydrolase in question to the *B. amyloliquefaciens* subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$R \text{ factor} = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

5

Equivalent residues which are functionally analogous to a specific residue of *B. amyloliquefaciens* subtilisin are defined as those amino acids of the precursor carbonyl hydrolases which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the *B. amyloliquefaciens* subtilisin as described herein. Further, they are those residues of the precursor carbonyl hydrolase (for which a tertiary structure has been obtained by x-ray crystallography), which occupy an analogous position to the extent that although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie within 0.13nm of the corresponding side chain atoms of *B. amyloliquefaciens* subtilisin. The three dimensional structures would be aligned as outlined above.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a mutant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally occurring sequence. The carbonyl hydrolase mutants of the present invention include the mature forms of carbonyl hydrolase mutants as well as the pro- and prepro-forms of such hydrolase mutants. The prepro-forms are the preferred construction since this facilitates the expression, secretion and maturation of the carbonyl hydrolase mutants.

"Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods disclosed in EPO Publication No. 0130756 to render them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing subtilisin is the *Bacillus* strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in EPO Publication No. 0130756 and further described by Yang, M.Y., et al. (1984) *J. Bacteriol.* 160, 15-21. Other host cells for expressing subtilisin include *Bacillus subtilis* 1168 (EPO Publication No. 0130756).

Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the carbonyl hydrolase mutants or expressing the desired carbonyl hydrolase mutant. In the case of vectors which encode the pre or prepro form of the carbonyl hydrolase mutant, such mutants, when expressed, are typically secreted from the host cell into the host cell medium.

"Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

The genes encoding the naturally-occurring precursor carbonyl hydrolase may be obtained in accord with the general methods described herein in EPO publication No. 0130756.

Once the carbonyl hydrolase gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-occurring precursor carbonyl hydrolase. Such modifications include the production of recombinant carbonyl hydrolases as disclosed in EPO

Publication No. 0130756 and the production of carbonyl hydrolase mutants described herein.

The carbonyl hydrolase mutants of the present invention may be generated by site specific mutagenesis (Smith, M. (1985) *Ann. Rev. Genet.* 423; Zoeller, M.J., et al. (1982) *Nucleic Acid Res.* 10, 6487-6500), cassette mutagenesis (EPO Publication No. 0130756) or random mutagenesis (Shortle, D., et al. (1985) *Genetics*, 110, 539; Shortle, D., et al. (1986) *Proteins: Structure, Function and Genetics*, 1, 81; Shortle, D. (1986) *J. Cell. Biochem.*, 30, 281; Alber, T., et al. (1985) *Proc. Natl. Acad. of Sci.*, 82, 747; Matsumura, M., et al. (1985) *J. Biochem.*, 260, 15298; Liao, H., et al. (1986) *Proc. Natl. Acad. of Sci.*, 83, 576) of the cloned precursor carbonyl hydrolase. Cassette mutagenesis and the random mutagenesis method disclosed herein are preferred.

The mutant carbonyl hydrolases expressed upon transformation of suitable hosts are screened for enzymes exhibiting one or more properties which are substantially different from the properties of the precursor carbonyl hydrolases, e.g., changes in substrate specificity, oxidative stability, thermal stability, alkaline stability, resistance to proteolytic degradation, pH-activity profiles and the like.

A change in substrate specificity is defined as a difference between the kcat/Km ratio of the precursor carbonyl hydrolase and that of the hydrolase mutant. The kcat/Km ratio is a measure of catalytic efficiency. Carbonyl hydrolase mutants with increased or diminished kcat/Km ratios are described in the examples. Generally, the objective will be to secure a mutant having a greater (numerically large) kcat/Km ratio for a given substrate, thereby enabling the use of the enzyme to more efficiently act on a target substrate. A substantial change in kcat/Km ratio is preferably at least 2-fold increase or decrease. However, smaller increases or decreases in the ratio (e.g., at least 1.5-fold) are also considered substantial. An increase in kcat/Km ratio for one substrate may be accompanied by a reduction in kcat/Km ratio for another substrate. This is a shift in substrate specificity, and mutants exhibiting such shifts have utility where the precursor hydrolase is undesirable, e.g. to prevent undesired hydrolysis of a particular substrate in an admixture of substrates. Km and kcat are measured in accord with known procedures, as described in EPO Publication No. 0130756 or as described herein.

Oxidative stability is measured either by known procedures or by the methods described hereinafter. A substantial change in oxidative stability is evidenced by at least about 50% increase or decrease (preferably decrease) in the rate of loss of enzyme activity when exposed to various oxidizing conditions. Such oxidizing conditions are exposure to the organic oxidant diperidodecanoic acid (DPDA) under the conditions described in the examples.

Alkaline stability is measured either by known procedures or by the methods described herein. A substantial change in alkaline stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half life of the enzymatic activity of a mutant when compared to the precursor carbonyl hydrolase. In the case of subtilisins, alkaline stability was measured as a function of autoprolytic degradation of subtilisin at alkaline pH, e.g. for example, 0.1M sodium phosphate, pH 12 at 25° or 30°C.

Thermal stability is measured either by known procedures or by the methods described herein. A substantial change in thermal stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half-life of the catalytic activity of a mutant when exposed to a relatively high temperature and neutral pH as compared to the precursor carbonyl hydrolase. In the case of subtilisins, thermal stability is measured by the autoprolytic degradation of subtilisin at elevated temperatures and neutral pH, e.g., for example 2mM calcium chloride, 50mM MOPS pH 7.0 at 59°C.

The inventors have produced mutant subtilisins containing the substitution of the amino acid residues of *B. amyloliquefaciens* subtilisin shown in Table I. The wild type amino acid sequence and DNA sequence of *B. amyloliquefaciens* subtilisin is shown in Fig. 1.

TABLE I

Residue	Replacement Amino Acid
Tyr21	F A
Thr22	C
Ser24	C
Asp32	Q S
Ser33	A T
Asp36	A G
Gly46	V
Ala48	E V R
Ser49	C L
Met50	C F V
Asn77	D
Ser87	C
Lys94	C
Val95	C
Leu96	D
Tyr104	A C D E F G H I K L M N P Q R S T V W
Ile107	V
Gly110	C R
Met124	I L
Asn155	A D H Q T
Glu156	Q S
Gly166	C E I L M P S T W Y
Gly169	C D E F H I K L M N P Q R T V W Y
Lys170	E R
Tyr171	F
Pro172	E Q
Phe189	A C D E G H I K L M N P Q R S T V W Y
Asp197	R A
Met199	I
Ser204	C R L P
Lys213	R T
Tyr217	A C D E F G H I K L M N P Q R S T V W
Ser221	A C

The different amino acids substituted are represented in Table I by the following single letter designations:

	Amino acid or residue thereof	3-letter symbol	1-letter symbol
	Alanine	Ala	A
	Glutamate	Glu	E
5	Glutamine	Gln	Q
	Aspartate	Asp	D
	Asparagine	Asn	N
	Leucine	Leu	L
	Glycine	Gly	G
10	Lysine	Lys	K
	Serine	Ser	S
	Valine	Val	V
	Arginine	Arg	R
	Threonine	Thr	T
15	Proline	Pro	P
	Isoleucine	Ile	I
	Methionine	Met	M
	Phenylalanine	Phe	F
	Tyrosine	Tyr	Y
20	Cysteine	Cys	C
	Tryptophan	Trp	W
	Histidine	His	H

25 Except where otherwise indicated by context, wild-type amino acids are represented by the above three-letter symbols and replaced amino acids by the above single-letter symbols. Thus, if the methionine at residue 50 in *B. amyloliquefaciens* subtilisin is replaced by phenylalanine, this mutation (mutant) may be designated Met50F or F50. Similar designations are used for multiple mutants.

30 In addition to the amino acids used to replace the residues disclosed in Table I, other replacements of amino acids at these residues are expected to produce mutant subtilisins having useful properties. These residues and replacement amino acids are shown in Table II.

35

40

45

50

55

TABLE II

Residue	Replacement Amino Acid(s)
Tyr-21	L
Thr22	K
Ser24	A
Asp32	
Ser33	G
Gly46	
Ala48	
Ser49	
Met50	L K I V
Asn77	D
Ser87	N
Lys94	R Q
Val95	L I
Tyr104	
Met124	K A
Ala152	C L I T M
Asn155	
Glu156	A T M L Y
Gly166	
Gly169	
Tyr171	K R E Q
Pro172	D N
Phe189	
Tyr217	
Ser221	
Met222	

Each of the mutant subtilisins in Table I contain the replacement of a single residue of the B. amyloliquefaciens amino acid sequence. These particular residues were chosen to probe the influence of such substitutions on various properties of B. amyloliquefacien subtilisin.

Thus, the inventors have identified Met124 and Met222 as important residues which if substituted with another amino acid produce a mutant subtilisin with enhanced oxidative stability. For Met124, Leu and Ile are preferred replacement amino acids. Preferred amino acids for replacement of Met222 are disclosed in EPO Publication No. 0130756.

Various other specific residues have also been identified as being important with regard to substrate specificity. These residues include Tyr104, Ala152, Glu156, Gly166, Gly169, Phe189 and Tyr217 for which mutants containing the various replacement amino acids presented in Table I have already been made, as well as other residues presented below for which mutants have yet to be made.

The identification of these residues, including those yet to be mutated, is based on the inventors' high resolution crystal structure of B. amyloliquefaciens subtilisin to 1.8 Å (see Table III), their experience with in vitro mutagenesis of subtilisin and the literature on subtilisin. This work and the x-ray crystal structures of subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449), product complexes (Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), and transition state analogs (Matthews, D.A., et al. (1975) J. Biol. Chem. 250, 7120-7126; Poulos, T.L., et al. (1976) J. Biol. Chem. 251, 1097-1103), has helped in identifying an extended peptide binding cleft in subtilisin. This substrate binding cleft together with substrate is schematically diagrammed in Fig. 2, according to the nomenclature of Schechter, I., et al. (1967) Biochem Bio. Res. Commun. 27, 157. The scissile bond in the substrate is identified by an arrow. The P and P' designations refer to the amino acids which are positioned respectively toward the amino or carboxy terminus relative to the scissile bond. The S and S' designations refer to subsites in the substrate binding cleft of subtilisin which interact with the corresponding substrate amino acid residues.

Atomic Coordinates for the
Apoenzyme Form of *B. Amylolyticus*
Subtilisin to 1.8Å Resolution

5

1	ALA N	19.434	53.195	-21.754	1	ALA CA	19.811	51.774	-21.965
1	ALA C	18.731	50.955	-21.324	1	ALA O	18.374	51.197	-20.175
1	ALA CO	21.099	51.518	-21.183	2	GLN N	18.248	49.886	-22.041
2	GLN CA	17.219	49.008	-21.434	2	GLN C	17.875	47.706	-20.992
2	GLN O	18.765	47.165	-21.691	2	GLN CO	16.125	48.760	-22.449
2	GLN CG	15.928	47.985	-21.921	2	GLN CD	13.912	47.762	-22.930
2	GLN OE1	13.023	48.612	-22.847	2	GLN NE2	14.115	46.917	-23.926
3	SER N	17.477	47.205	-19.852	3	SER CA	17.958	45.868	-19.437
3	SER C	16.735	44.918	-19.490	3	SER O	15.590	45.352	-19.229
3	SER CO	18.588	45.838	-18.869	3	SER OG	17.682	46.210	-17.049
4	VAL N	16.991	43.646	-19.725	4	VAL CA	15.946	42.619	-19.639
4	VAL C	16.129	41.934	-18.290	4	VAL O	17.123	41.178	-18.886
4	VAL CO	14.808	41.622	-20.822	4	VAL CG1	14.874	40.572	-20.741
4	VAL CG2	16.037	42.266	-22.186	5	PRO N	15.239	42.104	-17.331
5	PRO CA	15.384	41.415	-16.827	5	PRO C	15.501	39.905	-16.249
5	PRO O	14.885	39.263	-17.146	5	PRO CO	14.150	41.880	-15.263
5	PRO CG	13.841	43.215	-15.921	5	PRO CD	14.044	42.986	-17.417
6	TYR N	16.363	39.240	-15.487	6	TYR CA	16.628	37.883	-15.715
6	TYR C	15.359	36.975	-15.528	6	TYR O	15.224	35.943	-16.235
6	TYR CO	17.824	37.323	-14.834	6	TYR CG	18.021	35.847	-15.855
6	TYR CD1	18.437	35.452	-14.346	6	TYR CD2	17.696	34.988	-14.871
6	TYR CE1	18.535	34.870	-14.653	6	TYR CE2	17.815	33.539	-14.379
6	TYR C1	18.222	33.154	-15.628	6	TYR OH	18.312	31.838	-15.996
7	GLY N	14.464	37.362	-14.630	7	GLY CA	13.211	36.648	-14.376
7	GLY C	12.408	36.535	-15.670	7	GLY O	11.747	35.478	-15.883
8	VAL N	12.441	37.529	-16.541	8	VAL CA	11.777	37.523	-17.836
8	VAL C	12.363	36.433	-18.735	8	VAL O	11.639	35.716	-19.478
8	VAL CO	11.765	38.980	-18.567	8	VAL CG1	11.186	38.893	-19.943
8	VAL CG2	10.991	39.919	-17.733	9	SER N	13.661	36.318	-18.775
9	SER CA	14.419	35.342	-19.562	9	SER C	14.188	33.920	-18.965
9	SER O	14.112	33.814	-19.301	9	SER CO	15.926	35.632	-19.505
9	SER OG	16.162	36.747	-20.358	10	GLN N	14.115	33.887	-17.662
10	GLN CA	13.964	32.636	-16.876	10	GLN C	12.687	31.887	-17.277
10	GLN O	12.785	30.642	-17.413	10	GLN CO	14.125	32.885	-15.418
10	GLN CG	14.295	31.617	-14.588	10	GLN CD	14.486	31.911	-13.147
10	GLN OE1	14.554	33.068	-12.744	10	GLN NE2	14.552	30.969	-12.251
11	ILE N	11.625	32.575	-17.678	11	ILE CA	10.373	31.904	-18.182
11	ILE C	10.209	31.792	-19.605	11	ILE O	9.173	31.333	-20.188
11	ILE CO	9.132	32.669	-17.475	11	ILE CG1	9.066	34.117	-16.049
11	ILE CG2	9.162	32.655	-15.941	11	ILE CO1	7.588	34.648	-17.923
12	LYS N	11.272	32.185	-20.277	12	LYS CA	11.388	32.119	-21.722
12	LYS C	10.456	33.086	-22.522	12	LYS O	10.178	32.783	-23.686
12	LYS CO	11.257	30.646	-22.216	12	LYS CG	12.283	29.838	-21.423
12	LYS CD	12.543	28.517	-22.159	12	LYS CE	13.023	27.467	-21.166
12	LYS NE2	14.476	27.680	-20.935	13	ALA N	10.189	34.138	-21.991
13	ALA CA	9.325	35.198	-22.631	13	ALA C	10.026	35.716	-23.863
13	ALA O	9.338	35.804	-24.981	13	ALA CO	8.885	36.195	-21.565
14	PRO N	11.332	35.958	-23.893	14	PRO CA	11.985	36.430	-25.128
14	PRO C	11.786	35.957	-26.317	14	PRO O	11.778	36.847	-27.445
14	PRO CO	13.462	36.588	-24.492	14	PRO CG	13.328	36.978	-23.221
14	PRO CD	12.281	33.936	-22.758	15	ALA N	11.568	34.236	-26.119
15	ALA CA	11.379	33.458	-27.367	15	ALA C	10.882	33.795	-28.032
15	ALA O	10.888	33.718	-29.278	15	ALA CO	11.552	31.949	-27.062
16	LEU N	9.885	34.138	-27.248	16	LEU CA	7.791	34.558	-27.828
16	LEU C	7.912	35.925	-28.521	16	LEU O	7.342	36.126	-29.588
16	LEU CO	6.746	34.623	-26.698	16	LEU CG	5.798	33.465	-26.522
16	LEU CD1	5.881	33.234	-27.889	16	LEU CD2	6.694	32.287	-26.283
17	HIS N	8.665	36.828	-27.922	17	HIS CA	8.890	38.151	-28.538
17	HIS C	9.518	37.981	-29.898	17	HIS O	9.187	38.622	-30.856
17	HIS CO	9.788	39.188	-27.652	17	HIS CG	8.185	39.288	-26.262
17	HIS OD1	9.938	39.887	-25.272	17	HIS OD2	8.888	38.924	-25.674
17	HIS CE1	9.226	39.914	-24.144	17	HIS NE2	8.879	39.328	-24.381
18	SER N	10.443	37.833	-30.822	18	SER CA	11.189	36.738	-31.322

50

55

10	SLN C	30.139	30.123	-32.973	10	SLN D	30.047	30.112	-33.834
10	SLN CO	13.311	30.709	-31.172	10	SLN D2	13.321	30.480	-33.399
10	SLN M	9.090	30.405	-31.043	10	SLN CA	0.002	30.962	-32.878
10	SLN C	7.142	30.111	-33.303	10	SLN O	6.297	30.972	-34.219
10	SLN CO	7.221	30.049	-32.200	10	SLN CG	7.979	32.002	-31.823
10	SLN CD	6.023	31.707	-31.101	10	SLN OF1	0.719	31.033	-31.444
10	SLN M22	7.362	30.052	-30.256	20	SLN M	7.205	37.223	-32.907
20	SLV CA	6.369	30.307	-32.059	20	SLV C	5.101	30.492	-31.800
20	SLV O	6.263	30.276	-32.219	22	TVR M	5.202	37.801	-30.761
21	TVR CO	6.110	37.031	-29.763	21	TVR C	4.879	30.552	-28.923
21	TVR O	5.422	30.074	-27.756	21	TVR C1	3.498	36.431	-29.443
21	TVR CG	2.973	31.704	-30.700	21	TVR CD1	1.793	36.332	-31.238
21	TVR CD2	3.650	34.794	-31.997	21	TVR C21	1.306	33.797	-32.446
21	TVR C02	3.193	34.261	-32.888	21	TVR C2	2.003	34.798	-33.067
21	TVR OM	1.901	34.241	-34.250	22	TVR M	3.902	39.600	-28.208
22	TVR CA	6.262	40.927	-27.129	22	TVR O	3.091	40.922	-26.344
22	TVR O	3.297	41.723	-25.323	22	TVR CO	5.133	41.759	-27.611
22	TVR OG1	6.319	42.457	-28.997	22	TVR CG2	4.474	41.323	-28.229
23	GLV M	1.939	40.203	-26.493	23	GLV CA	0.009	40.400	-29.542
23	GLV C	-0.107	41.031	-26.118	23	GLV O	-1.013	42.095	-29.310
24	SLR M	-0.023	41.967	-27.371	24	SLR CO	-0.097	42.957	-28.012
24	SLR C	-2.303	42.426	-27.844	24	SLR O	-2.013	41.900	-28.160
24	SLR CO	-0.794	43.120	-29.920	24	SLR OG	0.063	43.632	-29.720
25	ASL M	-3.059	43.492	-27.513	25	ASL CA	-4.519	43.607	-27.303
25	ASL C	-9.013	42.073	-24.203	25	ASL O	-0.233	42.468	-26.190
25	ASL CO	-5.103	43.227	-28.703	25	ASL CG	-0.960	44.170	-29.005
25	ASL OD1	-4.943	43.747	-31.003	25	ASL OD2	-4.747	45.461	-29.904
26	VAL M	-4.177	42.449	-25.292	26	VAL CA	-4.674	41.679	-24.143
26	VAL C	-4.792	42.652	-22.957	26	VAL O	-3.050	43.419	-22.609
26	VAL CO	-3.714	40.903	-23.821	26	VAL CG1	-4.160	39.802	-22.940
26	VAL CG2	-3.998	39.576	-25.018	27	LVS M	-5.910	42.613	-22.301
27	LVS CA	-0.133	43.524	-21.175	27	LVS C	-5.015	42.872	-19.041
27	LVS O	-6.403	41.073	-19.433	27	LVS CO	-7.590	43.901	-21.149
27	LVS CG	-0.046	44.875	-22.490	27	LVS CD	-9.321	45.302	-22.020
27	LVS C0	-10.304	40.497	-23.137	27	LVS M2	-0.686	46.253	-24.264
28	VAL M	-4.819	43.462	-19.200	28	VAL CA	-4.497	42.930	-17.097
28	VAL C	-4.759	43.959	-16.928	28	VAL O	-4.209	45.095	-16.017
28	VAL CO	-2.924	42.666	-17.932	28	VAL CG1	-2.466	42.103	-16.509
28	VAL CG2	-2.667	41.003	-19.173	29	ALA M	-5.484	43.327	-15.013
29	ALA CA	-5.747	44.330	-14.639	29	ALA C	-4.750	44.010	-15.503
29	ALA O	-4.666	42.043	-13.104	29	ALA CO	-7.172	44.107	-14.101
30	VAL M	-4.057	40.033	-13.072	30	VAL CA	-3.146	44.962	-11.910
30	VAL C	-3.998	43.409	-10.601	30	VAL O	-4.199	46.640	-10.970
30	VAL CO	-1.006	45.010	-12.149	30	VAL CG1	-0.996	40.901	-10.900
30	VAL CG2	-1.033	45.236	-13.307	31	ILR M	-4.514	44.510	-9.077
31	ILR CA	-5.320	44.046	-8.679	31	ILR C	-4.346	44.933	-7.866
31	ILR O	-3.029	43.913	-6.097	31	ILR CO	-6.457	43.774	-8.001
31	ILR CG1	-7.208	43.707	-9.799	31	ILR CG2	-7.270	44.038	-7.223
31	ILR CD1	-8.617	42.056	-9.717	32	ASP M	-4.044	46.193	-7.227
32	ASP CA	-2.964	46.467	-6.253	32	ASP C	-3.071	47.009	-5.703
32	ASP O	-6.197	48.418	-5.302	32	ASP CO	-1.693	46.129	-7.097
32	ASP CG	-0.453	45.702	-6.273	32	ASP OD1	0.034	44.592	-6.576
32	ASP OD2	-0.001	46.420	-5.330	33	SRM M	-1.921	40.512	-3.394
33	SRM CA	-1.093	49.037	-4.001	33	SRM C	-1.902	40.976	-3.038
33	SRM O	-1.704	32.136	-5.363	33	SRM CO	-0.621	49.922	-3.939
33	SRM OG	0.533	50.025	-4.774	34	SLV M	-2.173	50.769	-7.004
34	GLV CA	-2.255	51.728	-5.163	34	GLV C	-1.039	51.648	-9.057
34	GLV O	-0.144	50.031	-5.761	35	ILR M	-0.965	52.431	-10.102
35	ILR CA	0.208	52.430	-10.093	35	ILR C	0.368	53.919	-11.263
35	ILR O	-0.327	54.630	-11.744	35	ILR CO	-0.042	51.694	-12.367
35	ILR CG1	-0.530	50.210	-12.097	35	ILR CG2	3.149	51.741	-13.362
40	ILR CD1	-0.962	49.493	-13.424	36	ASP M	1.016	54.253	-10.971
36	ASP CA	2.350	55.630	-11.232	36	ASP C	2.203	55.956	-12.702

45

50

55

EP 0 251 446 B1

	36	ASP D	3.004	55.471	-13.579	36	ASP CB	3.712	55.720	-10.514
	36	ASP CG	4.339	57.099	-10.804	36	ASP OD1	3.755	57.974	-11.429
	36	ASP OD2	3.440	57.277	-10.263	37	SER H	1.304	56.822	-13.111
	37	SER CA	3.183	57.221	-10.512	37	SER C	2.377	58.095	-14.949
	37	SER D	2.545	58.303	-10.151	37	SER CB	-0.093	58.049	-14.788
	37	SER DG	-0.070	59.133	-13.079	38	SER H	3.163	58.614	-14.001
5	38	SER CA	4.261	59.505	-14.407	38	SER C	5.464	58.705	-14.992
	38	SER D	0.543	59.251	-15.205	38	SER CB	4.742	60.435	-13.399
	38	SER DG	9.316	59.865	-12.234	39	NIS H	5.454	57.390	-14.892
	39	NIS CA	6.637	56.574	-15.291	39	NIS C	6.601	56.401	-16.778
	39	NIS D	5.738	55.878	-17.419	39	NIS CB	6.637	55.203	-14.515
	39	NIS CG	0.014	54.609	-14.456	39	NIS OD1	0.795	54.354	-15.561
	39	NIS OD2	0.749	54.345	-13.389	39	NIS CB1	9.970	53.930	-15.130
10	39	NIS HE2	9.986	53.910	-13.808	40	PRO H	7.007	54.834	-17.387
	40	PRO CA	7.988	56.697	-18.831	40	PRO C	0.154	55.280	-19.357
	40	PRO D	0.032	55.097	-20.578	40	PRO CB	9.247	57.533	-19.161
	40	PRO CG	10.053	57.405	-17.902	40	PRO CD	0.988	57.452	-16.776
	41	ASP H	0.461	54.320	-18.485	41	ASP OD2	11.140	58.399	-18.668
	41	ASP OD1	10.325	51.395	-20.429	41	ASP CG	10.473	51.387	-19.211
	41	ASP CB	9.799	52.239	-18.224	41	ASP CA	0.445	52.959	-18.964
	41	ASP C	7.311	52.163	-18.839	41	ASP D	7.396	50.947	-18.977
15	42	LEU H	4.185	52.003	-18.558	42	LEU CA	4.892	52.147	-18.466
	42	LEU C	3.924	52.907	-19.376	42	LEU D	3.993	54.163	-19.490
	42	LEU CB	4.421	52.158	-17.008	42	LEU CG	5.102	51.363	-15.946
	42	LEU CD1	4.535	51.546	-14.581	42	LEU CD2	5.273	49.877	-16.358
	43	LVS H	3.010	52.135	-19.946	43	LVS CA	1.893	52.485	-20.721
	43	LVS C	0.637	52.156	-20.018	43	LVS D	0.504	50.920	-19.820
	43	LVS CB	2.021	52.389	-22.169	43	LVS CG	0.685	52.436	-22.910
	43	LVS CD	0.998	52.862	-24.339	43	LVS CE	-0.180	52.584	-25.260
20	43	LVS M2	0.337	51.757	-26.418	44	VAL H	-0.191	53.035	-19.490
	44	VAL CA	-1.407	52.039	-19.765	44	VAL C	-2.571	52.087	-19.731
	44	VAL D	-2.623	53.906	-20.434	44	VAL CB	-1.480	53.351	-17.383
	44	VAL CG1	-2.724	52.941	-16.502	44	VAL CG2	-0.197	53.194	-16.553
	45	ALA H	-3.494	51.951	-19.071	45	ALA CA	-4.619	51.977	-20.810
	45	ALA C	-5.841	52.507	-20.053	45	ALA D	-6.703	53.085	-20.703
	45	ALA CB	-4.031	50.500	-21.389	46	GLY H	-5.910	52.354	-18.748
	46	GLY CA	-7.082	52.037	-18.081	46	GLY C	-6.987	52.443	-14.538
25	46	GLY D	-5.938	52.006	-16.035	47	GLY H	-8.092	52.650	-15.793
	47	GLY CA	-8.014	52.246	-14.388	47	GLY C	-9.179	52.757	-13.572
	47	GLY D	-9.988	53.481	-14.185	48	ALA H	-9.221	52.446	-12.330
	48	ALA CA	-10.255	52.878	-11.382	48	ALA C	-9.790	52.675	-9.968
	48	ALA D	-9.066	51.720	-9.725	48	ALA CB	-11.550	52.100	-11.617
	49	SER H	-10.149	53.547	-9.037	49	SER CA	-9.752	53.355	-7.652
	49	SER C	-10.947	52.984	-6.783	49	SER D	-11.972	53.677	-6.908
30	49	SER CB	-9.092	54.588	-7.029	49	SER DG	-8.879	54.255	-5.650
	50	RET H	-10.835	52.007	-5.932	50	RET CA	-11.052	51.549	-4.974
	50	RET C	-11.463	51.962	-3.561	50	RET D	-11.997	51.398	-2.575
	50	RET CB	-12.012	50.018	-4.996	50	RET CG	-11.912	49.463	-6.389
	50	RET SD	-13.468	49.889	-7.254	50	RET CE	-12.000	50.111	-0.903
	51	VAL H	-10.427	52.760	-3.422	51	VAL CA	-9.968	53.170	-2.067
	51	VAL C	-10.630	54.562	-1.987	51	VAL D	-10.237	55.437	-2.682
	51	VAL CB	-0.443	53.155	-2.000	51	VAL CG1	-7.092	53.579	-0.631
35	51	VAL CG2	-7.764	53.015	-2.302	52	PRO H	-11.621	54.693	-1.056
	52	PRO CA	-12.372	55.933	-0.821	52	PRO C	-11.490	57.123	-0.440
	52	PRO D	-11.771	50.220	-0.925	52	PRO CB	-13.400	55.994	0.244
	52	PRO CG	-13.583	54.103	0.085	52	PRO CD	-17.264	53.620	-0.175
	53	SER H	-10.442	56.984	0.299	53	SER CA	-9.538	57.982	0.682
	53	SER C	-8.428	58.245	-0.326	53	SER D	-7.479	59.224	-0.038
	53	SER CB	-9.004	57.707	2.069	53	SER DG	-8.256	56.521	2.127
	54	GLU H	-8.234	57.523	-1.393	54	GLU CA	-7.804	57.048	-2.421
	54	GLU C	-7.767	57.383	-3.785	54	GLU D	-7.533	56.243	-4.379
40	54	GLU CB	-8.134	56.599	-2.154	54	GLU CG	-5.289	56.959	-0.927
	54	GLU CD	-6.044	54.049	-0.078	54	GLU OD1	-5.045	55.604	-1.968

45

50

55

54	GLN DE2	-9.900	55.777	0.271	55	TMR H	-0.971	58.251	-4.249
55	TMR CA	-9.433	58.121	-5.441	55	TMR C	-0.764	58.139	-6.779
55	TMR O	-9.433	57.919	-7.010	55	TMR CB	-10.906	59.200	-5.303
55	TMR CG1	-9.885	60.310	-5.418	55	TMR CG2	-11.432	59.143	-4.811
56	ASN H	-7.482	58.403	-6.877	56	ASN WD2	-6.930	61.179	-9.881
56	ASN OD1	-5.875	58.967	-10.337	56	ASN CG	-5.273	59.925	-9.555
56	ASN CB	-5.898	59.694	-8.208	56	ASN CA	-6.762	58.425	-8.200
56	ASN C	-6.812	57.094	-8.305	56	ASN O	-5.104	56.046	-7.478
57	PRO H	-6.362	56.261	-9.258	57	PRO CG	-7.123	55.257	-11.177
57	PRO CD	-7.384	56.433	-10.272	57	PRO CB	-6.444	54.178	-10.235
57	PRO CA	-5.679	54.961	-9.332	57	PRO C	-4.301	55.082	-9.966
57	PRO O	-3.589	54.128	-9.945	58	PME H	-3.998	56.262	-10.491
58	PME CA	-2.747	56.577	-11.222	58	PME C	-1.712	57.129	-10.253
58	PME O	-0.635	57.497	-10.600	58	PME CB	-2.943	57.582	-12.423
58	PME CG	-3.983	56.968	-13.337	58	PME CD1	-3.756	55.780	-14.059
58	PME CD2	-5.211	57.630	-13.459	58	PME CE1	-6.722	55.255	-14.928
58	PME CE2	-6.194	57.095	-14.276	58	PME C2	-5.949	55.939	-15.051
59	GLN H	-2.064	57.119	-8.990	59	GLN CA	-1.172	57.583	-7.934
59	GLN C	-0.887	56.403	-7.800	59	GLN O	-1.639	56.083	-6.115
59	GLN CB	-1.862	58.668	-7.089	59	GLN CG	-0.942	59.261	-6.834
59	GLN CD	-1.790	60.157	-5.150	59	GLN DE1	-1.484	61.288	-4.836
59	GLN DE2	-2.959	59.685	-6.742	60	ASP H	0.410	55.895	-7.211
60	ASP CA	0.851	54.792	-6.304	60	ASP C	1.631	55.267	-5.090
60	ASP O	2.827	55.550	-5.231	60	ASP CB	1.396	53.744	-7.188
60	ASP CG	2.077	52.538	-6.380	60	ASP OD1	1.746	52.337	-5.190
60	ASP OD2	2.915	51.841	-7.030	61	ASN H	0.959	55.265	-3.950
61	ASN WD2	-1.364	57.747	-2.347	61	ASN BD1	0.666	58.566	-2.875
61	ASN CG	-0.040	57.670	-2.399	61	ASN CB	0.531	56.401	-1.784
61	ASN CA	1.557	55.734	-2.700	61	ASN C	2.291	54.632	-1.940
61	ASN O	2.933	54.862	-0.902	62	ASN H	2.210	53.434	-2.468
62	ASN CA	2.877	52.348	-1.709	62	ASN C	4.124	51.893	-2.479
62	ASN O	4.951	51.313	-3.770	62	ASN CB	1.783	51.319	-1.621
62	ASN CG	2.371	50.103	-0.697	62	ASN OD1	2.633	49.077	-1.343
62	ASN WD2	2.622	50.208	0.601	63	SER H	4.152	52.104	-3.761
63	SER CA	5.189	51.696	-4.709	63	SER C	5.071	50.256	-5.209
63	SER O	5.593	49.790	-6.269	63	SER CB	6.523	51.958	-4.812
63	SER CG	6.871	50.698	-3.418	64	MIS H	4.202	49.475	-4.639
64	MIS CA	3.994	68.856	-4.935	64	MIS C	3.366	47.759	-6.261
64	MIS O	3.861	46.974	-7.108	64	MIS CB	3.184	47.501	-3.747
64	MIS CG	3.144	46.021	-3.726	64	MIS WD1	2.107	45.247	-4.241
64	MIS CD2	4.054	45.194	-3.135	64	MIS CE1	2.416	43.966	-4.054
64	MIS ME2	3.556	43.920	-3.368	65	GLY H	2.287	48.428	-6.587
65	GLY CA	1.552	48.264	-7.830	65	GLY C	2.392	48.636	-9.037
65	GLY O	2.238	48.078	-10.134	66	TMR H	3.233	49.659	-8.832
66	TMR CA	4.064	50.117	-9.954	66	TMR C	5.089	49.009	-10.291
66	TMR O	5.333	48.789	-11.463	66	TMR CB	4.744	51.511	-9.667
66	TMR CG1	3.637	52.425	-9.406	66	TMR CG2	5.536	52.078	-10.849
67	MIS H	5.685	48.443	-9.274	67	MIS CA	6.703	47.341	-9.458
67	MIS C	6.091	46.141	-10.143	67	MIS O	6.649	45.630	-11.150
67	MIS CB	7.300	47.871	-8.064	67	MIS CG	8.395	46.275	-8.148
67	MIS CD1	8.590	44.987	-8.276	67	MIS CD2	9.904	46.678	-8.876
67	MIS CE1	9.857	44.691	-8.299	67	MIS ME2	10.678	45.514	-8.186
68	VAL O	4.892	45.749	-9.731	68	VAL CA	4.142	44.607	-10.266
68	VAL C	3.856	44.460	-11.740	68	VAL O	4.114	43.942	-12.535
68	VAL CB	2.939	44.252	-9.386	68	VAL CG1	1.960	43.260	-10.820
68	VAL CG2	3.319	43.785	-8.888	69	ALA H	3.373	46.049	-12.113
69	ALA CA	3.037	46.446	-13.429	69	ALA C	4.193	46.390	-14.411
69	ALA O	4.028	45.913	-13.565	69	ALA CB	2.332	47.851	-13.386
70	GLY H	5.348	46.782	-13.914	70	GLY CA	6.595	46.805	-14.670
70	GLY C	7.046	45.378	-15.021	70	GLY O	7.604	45.154	-16.119
71	TMR H	6.820	44.431	-14.138	71	TMR CA	7.177	43.019	-16.444
71	TMR C	6.224	42.506	-15.543	71	TMR O	6.682	41.828	-16.695
71	TMR CB	7.119	42.878	-13.191	71	TMR CG1	8.191	42.592	-12.390

45

50

55

EP 0 251 446 B1

	71	YMR CG2	7.274	40.903	-13.596	72	VAL M	6.930	62.007	-13.627
	72	VAL CA	3.976	42.491	-16.604	72	VAL C	6.312	63.004	-17.031
	72	VAL O	4.341	42.300	-10.860	71	VAL CO	2.916	62.067	-16.085
	72	VAL CG1	1.512	42.400	-17.170	72	VAL CG2	2.142	62.327	-16.723
5	73	ALA M	4.504	44.437	-17.900	73	ALA CA	4.907	63.001	-19.167
	73	ALA C	5.433	46.333	-19.355	73	ALA O	5.062	67.100	-20.216
	73	ALA CO	3.107	45.441	-19.433	76	ALA M	6.544	66.429	-16.635
	74	ALA CA	7.470	47.591	-18.959	74	ALA C	7.740	67.640	-20.342
	74	ALA O	7.959	46.640	-21.054	76	ALA CO	8.633	67.444	-17.925
	75	LEU M	7.650	48.704	-21.039	75	LEU CA	7.012	68.968	-22.456
	75	LEU C	9.192	40.568	-22.966	75	LEU O	10.142	68.750	-22.253
	75	LEU CO	7.540	50.471	-22.009	75	LEU CG	6.123	50.913	-22.379
10	75	LEU CD1	6.079	52.436	-22.300	75	LEU CD2	5.094	50.442	-21.405
	76	ASM M	9.147	40.103	-24.169	76	ASM MD2	12.305	46.432	-26.304
	76	ASM BD1	10.950	45.040	-27.928	76	ASM CG	11.195	46.274	-26.002
	76	ASM CO	10.010	46.651	-25.908	76	ASM CA	10.359	47.730	-26.930
	76	ASM C	10.703	49.040	-25.643	76	ASM O	10.157	49.479	-26.619
	77	ASM M	11.004	49.464	-25.071	77	ASM CA	12.220	50.957	-25.601
	77	ASM C	13.707	51.029	-25.348	77	ASM O	14.364	49.979	-25.333
	77	ASM CO	11.335	52.076	-25.117	77	ASM CG	11.250	52.027	-23.616
15	77	ASM CD1	12.032	51.346	-22.917	77	ASM MD2	10.294	52.741	-23.025
	70	SER M	14.125	52.267	-25.164	78	SER CA	15.513	52.614	-26.906
	70	SER C	15.010	52.742	-23.436	78	SER O	16.902	53.071	-23.164
	70	SER CO	15.905	53.941	-25.507	78	SER CG	15.926	53.070	-26.999
	79	ILE M	14.050	52.565	-22.529	79	ILE CA	15.155	52.704	-21.120
	79	ILE C	14.617	51.603	-20.230	79	ILE O	13.043	50.041	-20.679
	79	ILE CO	14.471	54.174	-20.697	79	ILE CG1	12.949	54.032	-20.014
	79	ILE CG2	14.997	55.320	-21.612	79	ILE CD1	12.135	55.176	-20.155
20	00	GLY M	14.995	51.740	-18.901	00	GLY CA	14.476	50.940	-17.913
	00	GLY C	14.612	49.448	-10.219	00	GLY O	15.719	48.994	-18.944
	01	VAL M	13.513	40.766	-17.980	01	VAL CA	13.411	47.206	-18.061
	01	VAL C	12.511	46.919	-19.217	01	VAL O	12.260	47.739	-20.117
	01	VAL CO	13.001	46.755	-16.677	01	VAL CG1	14.030	47.004	-15.573
	01	VAL CG2	11.630	47.261	-16.231	02	LEU M	12.126	45.645	-19.216
	02	LEU CA	11.312	45.020	-20.256	02	LEU C	10.390	44.020	-19.510
	02	LEU O	10.050	43.356	-10.600	02	LEU CO	12.206	44.219	-21.229
25	02	LEU CG	11.430	43.568	-22.366	02	LEU CD1	10.796	44.657	-23.223
	02	LEU CD2	12.359	42.675	-23.192	03	GLY M	9.131	44.100	-19.016
	03	GLY CA	8.133	43.321	-19.114	03	GLY C	0.027	42.011	-19.925
	03	GLY O	0.946	41.022	-21.026	04	VAL M	7.272	41.112	-19.203
	04	VAL CA	6.973	39.007	-19.088	04	VAL C	6.164	40.030	-21.140
	04	VAL O	6.424	39.472	-22.194	04	VAL CO	6.256	38.920	-10.041
	04	VAL CG1	5.680	37.677	-19.557	04	VAL CG2	7.190	38.507	-17.705
30	05	ALA M	5.156	40.924	-21.024	05	ALA CA	4.217	41.194	-22.150
	05	ALA C	4.213	42.603	-22.396	05	ALA O	3.260	43.401	-22.030
	05	ALA CO	2.046	40.663	-21.748	06	PRO M	5.240	43.106	-23.059
	06	PRO CA	5.413	44.635	-23.205	06	PRO C	4.321	45.371	-23.947
	06	PRO CO	4.291	46.805	-23.049	06	PRO CO	4.522	44.704	-23.013
	06	PRO CD	7.030	43.466	-24.546	06	PRO CD	6.377	62.440	-23.636
	07	SER M	3.540	44.676	-24.769	07	SER CA	2.409	45.324	-25.529
	07	SER C	1.103	45.132	-24.097	07	SER O	0.162	45.513	-25.619
	07	SER CO	2.401	44.777	-26.927	07	SER CG	3.591	45.143	-27.503
35	08	ALA M	1.017	44.504	-23.742	08	ALA CO	-0.163	43.510	-21.020
	08	ALA CA	-0.273	44.353	-23.004	08	ALA C	-0.090	45.717	-22.690
	08	ALA O	-0.174	46.717	-22.435	09	SER M	-2.219	45.691	-22.678
	09	SER CO	-4.146	47.102	-24.280	09	SER CA	-4.343	46.903	-22.090
	09	SER CA	-3.001	46.067	-22.227	09	SER C	-3.136	46.700	-20.727
	09	SER O	-3.793	45.064	-20.209	10	LEU M	-2.446	47.656	-20.037
	09	LEU CA	-2.370	47.667	-10.593	10	LEU C	-1.403	48.430	-17.064
40	09	LEU O	-3.502	49.604	-10.215	10	LEU CO	-0.951	40.273	-10.426
	09	LEU CG	-0.233	47.851	-17.174	10	LEU CD1	-0.076	46.341	-17.219
	09	LEU CD2	-1.160	40.524	-17.047	11	TYR M	-4.264	47.964	-16.930
	09	TYR CA	-5.258	40.670	-16.137	11	TYR C	-4.073	48.750	-16.605

45

50

55

5	91	TYR D	-6.494	47.749	-14.923	91	TYR C8	-6.686	48.093	-14.314
	91	TYR C6	-7.894	48.237	-17.741	91	TYR CD1	-6.595	47.415	-10.755
	91	TYR CD2	-7.971	49.275	-18.149	91	TYR CE1	-6.985	47.572	-20.098
	91	TYR CE2	-8.315	49.421	-19.492	91	TYR C2	-7.794	48.362	-20.463
	91	TYR DM	-8.102	48.752	-21.764	92	ALA M	-6.895	49.958	-14.104
	92	ALA CA	-4.949	50.199	-12.787	92	ALA C	-5.823	50.833	-11.903
	92	ALA O	-6.723	50.898	-12.850	92	ALA C8	-3.997	51.621	-12.488
	93	VAL M	-3.959	48.993	-31.329	93	VAL CA	-7.183	48.854	-10.325
	93	VAL C	-6.708	49.814	-8.899	93	VAL O	-6.181	47.993	-8.372
	93	VAL C8	-7.957	47.555	-10.633	93	VAL CG1	-9.213	47.488	-9.725
	93	VAL CG2	-8.195	47.378	-12.872	94	LVS M	-6.987	50.217	-8.227
	94	LVS CA	-6.378	50.464	-6.999	94	LVS C	-7.351	49.985	-5.894
10	94	LVS O	-8.458	50.480	-5.783	94	LVS C8	-6.851	51.976	-6.818
	94	LVS CG	-5.394	52.320	-5.467	94	LVS CD	-4.868	53.785	-5.582
	94	LVS CE	-4.399	54.208	-4.199	94	LVS M2	-3.735	55.544	-4.387
	95	VAL M	-4.909	49.071	-5.026	95	VAL CA	-7.646	48.457	-3.920
	95	VAL C	-6.919	48.499	-2.568	95	VAL O	-7.425	48.156	-1.581
	95	VAL C8	-8.184	47.038	-4.319	95	VAL CG1	-8.868	46.852	-5.619
	95	VAL CG2	-6.980	46.180	-4.332	96	LEU M	-5.676	48.974	-2.604
	96	LEU CA	-4.782	49.193	-1.886	96	LEU C	-4.331	50.559	-1.321
15	96	LEU O	-3.942	51.121	-2.336	96	LEU C8	-3.589	48.241	-1.573
	96	LEU CG	-3.593	46.799	-2.072	96	LEU CD1	-2.287	46.184	-2.163
	96	LEU CD2	-4.489	46.882	-1.845	97	GLY M	-4.326	50.975	-0.886
	97	GLY CA	-3.890	52.307	0.287	97	GLY C	-2.363	52.437	0.385
	97	GLY O	-2.619	51.463	0.165	98	ALA M	-1.954	53.640	0.758
	98	ALA C8	-0.428	55.478	1.510	98	ALA CA	-8.563	54.868	0.965
	98	ALA C	0.188	53.118	1.917	98	ALA O	1.393	52.921	1.663
	99	ASP M	-0.504	52.573	2.912	99	ASP CD2	-2.631	51.842	6.151
20	99	ASP CD1	-2.730	50.982	4.883	99	ASP C8	-2.883	51.131	5.040
	99	ASP C8	-8.648	51.603	5.175	99	ASP CA	0.181	51.618	3.855
	99	ASP C	0.146	50.165	3.320	99	ASP O	0.735	49.313	4.029
	100	GLY M	-8.424	49.883	2.168	100	GLY CA	-0.343	48.521	3.615
	100	GLY C	-1.520	47.651	2.802	100	GLY O	-3.649	46.512	3.479
	101	SER M	-2.342	48.128	2.908	101	SER CA	-3.542	47.388	3.315
	101	SER C	-4.759	47.894	2.532	101	SER O	-4.758	48.972	1.907
25	101	SER C8	-3.716	47.447	4.817	101	SER CG	-4.411	48.634	5.209
	102	GLY M	-5.821	47.892	2.577	102	GLY CA	-7.877	47.422	1.896
	102	GLY C	-8.166	46.536	2.528	102	GLY O	-7.888	45.631	3.830
	103	GLM M	-9.377	47.058	2.498	103	GLM C8	-10.535	46.297	3.820
	103	GLM C	-10.963	45.232	2.022	103	GLM	-10.779	45.482	0.817
	103	GLM C8	-11.671	47.387	3.274	103	GLM CG	-11.368	48.805	4.586
	103	GLM CD	-12.368	49.184	4.915	103	GLM DE1	-12.159	49.816	5.902
	103	GLM ME2	-13.619	49.197	4.112	104	TYR M	-11.611	44.141	2.451
30	104	TYR CA	-12.868	43.126	1.588	104	TYR C	-13.031	43.690	0.473
	104	TYR O	-12.939	43.276	-0.887	104	TYR C8	-12.697	41.866	2.143
	104	TYR CG	-11.629	48.829	2.472	104	TYR CD1	-11.819	39.789	3.377
	104	TYR CD2	-10.379	48.959	1.840	104	TYR CE1	-10.809	38.885	3.707
	104	TYR CE2	-9.352	48.057	2.371	104	TYR C2	-9.564	39.822	3.081
	104	TYR DM	-8.481	38.191	3.324	105	SER M	-13.909	44.572	0.903
	105	SER CA	-14.877	45.166	-0.034	105	SER C	-14.172	45.920	-1.159
35	105	SER O	-14.759	45.935	-2.258	105	SER C8	-15.880	46.121	0.681
	105	SER CG	-15.289	47.839	1.450	106	TRP M	-13.879	46.625	-0.834
	106	TRP CA	-12.421	47.391	-1.948	106	TRP C	-11.895	46.436	-3.012
	106	TRP O	-12.821	46.648	-4.245	106	TRP C8	-11.321	48.254	-1.355
	106	TRP CG	-11.645	48.111	-0.286	106	TRP CD1	-12.862	49.524	0.264
	106	TRP CD2	-10.658	49.812	0.981	106	TRP ME1	-12.691	50.358	1.360
	106	TRP CE2	-11.359	50.573	1.561	106	TRP CE3	-9.275	49.852	0.576
	106	TRP CE2	-10.671	51.318	2.500	106	TRP CE3	-8.568	50.563	1.525
40	106	TRP CE2	-9.293	51.291	2.455	107	ILE M	-11.339	45.330	-2.481
	107	ILE CA	-10.765	44.250	-3.375	107	ILE C	-11.955	43.594	-4.190
	107	ILE O	-11.695	43.674	-5.398	107	ILE C8	-9.966	43.183	-2.523
	107	ILE CG1	-8.634	43.784	-1.974	107	ILE CG2	-9.632	41.930	-3.381
	107	ILE CD1	-8.263	42.998	-0.627	108	ILE M	-12.994	43.292	-3.577

45

50

55

EP 0 251 446 B1

	100	ILE CA	-14.116	42.722	-4.373	108	ILE C	-14.639	43.694	-5.386
	100	ILE O	-14.894	43.329	-6.552	108	ILE CO	-15.246	42.263	-3.320
	100	ILE CG1	-14.726	41.077	-2.482	108	ILE CG2	-16.568	42.024	-4.093
	100	ILE CD1	-15.452	40.865	-1.331	109	ASN M	-14.751	44.958	-4.981
	109	ASN CA	-15.204	46.010	-5.916	109	ASN C	-14.232	46.067	-7.004
	109	ASN O	-14.668	46.272	-0.235	109	ASN CO	-15.200	47.359	-5.207
5	109	ASN CG	-16.528	47.486	-4.353	109	ASN CD1	-17.655	46.695	-4.646
	109	ASN CD2	-16.633	48.467	-3.442	110	GLY M	-12.951	45.908	-6.774
	110	GLY CA	-11.952	45.917	-7.865	110	GLY C	-12.108	44.712	-8.012
	110	GLY O	-11.929	44.929	-10.034	111	ILE M	-12.379	43.539	-8.246
	111	ILE CA	-12.603	42.334	-9.099	111	ILE C	-13.059	42.560	-9.942
	111	ILE O	-12.921	42.304	-11.148	111	ILE CO	-12.734	48.948	-8.366
	111	ILE CG1	-12.421	40.501	-7.655	111	ILE CG2	-13.122	39.791	-9.347
	111	ILE CD1	-12.508	39.706	-6.336	112	GLU M	-14.893	43.875	-9.280
10	112	GLU CA	-16.118	43.376	-10.066	112	GLU C	-15.872	44.347	-11.171
	112	GLU O	-16.467	44.130	-12.246	112	GLU CO	-17.229	43.899	-9.141
	112	GLU CG	-17.067	42.917	-8.135	112	GLU CD	-18.724	41.824	-8.685
	112	GLU CD1	-19.061	40.866	-0.816	112	GLU CD2	-19.123	41.928	-9.866
	113	TRP M	-15.094	45.403	-10.971	113	TRP CA	-14.756	46.400	-12.000
	113	TRP C	-14.876	45.663	-13.140	113	TRP O	-14.319	45.932	-14.332
	113	TRP CO	-13.882	47.553	-11.434	113	TRP CG	-13.486	48.356	-12.481
	113	TRP CD1	-16.148	49.736	-12.681	113	TRP CD2	-12.441	48.952	-13.463
15	113	TRP CD2	-13.597	50.443	-13.723	113	TRP CD3	-12.545	49.761	-14.215
	113	TRP CD3	-11.451	47.645	-13.809	113	TRP CD4	-11.696	50.045	-15.274
	113	TRP CD5	-10.610	47.899	-14.879	113	TRP CD6	-10.752	49.874	-15.603
	114	ALA M	-13.089	44.801	-12.832	114	ALA CA	-12.333	44.865	-13.874
	114	ALA C	-13.199	43.179	-14.752	114	ALA O	-12.963	43.874	-15.978
	114	ALA CO	-11.299	43.192	-13.140	115	ILE M	-14.174	42.340	-14.119
	115	ILE CA	-15.070	41.640	-14.097	115	ILE C	-15.928	42.485	-15.856
20	115	ILE O	-16.077	42.225	-17.070	115	ILE CO	-16.000	40.840	-13.922
	115	ILE CG1	-15.218	39.836	-13.843	115	ILE CG2	-17.151	40.168	-14.755
	115	ILE CD1	-16.004	39.411	-11.763	116	ALA M	-16.534	43.527	-15.267
	116	ALA CA	-17.390	44.440	-16.850	116	ALA C	-16.786	45.069	-17.278
	116	ALA O	-17.323	45.255	-18.363	116	ALA CO	-18.011	45.510	-15.151
	117	ASN M	-15.423	45.390	-17.122	117	ASN CA	-14.553	45.967	-18.139
	117	ASN C	-13.827	44.974	-19.034	117	ASN O	-12.997	45.436	-19.820
	117	ASN CO	-13.615	44.958	-17.426	117	ASN CG	-14.400	48.177	-16.939
25	117	ASN CD1	-14.565	49.882	-17.773	117	ASN CD2	-14.931	48.249	-15.736
	118	ASN M	-14.223	43.723	-18.967	118	ASN CA	-13.760	42.642	-19.832
	118	ASN C	-12.240	42.444	-19.843	118	ASN CO	-11.617	42.309	-20.932
	118	ASN CO	-14.247	42.863	-21.279	118	ASN CG	-15.737	43.060	-21.395
	118	ASN CD1	-16.910	42.321	-20.759	118	ASN CD2	-16.136	44.096	-22.133
	119	MET M	-11.686	42.500	-18.675	119	MET CA	-10.232	42.222	-18.478
	119	MET C	-10.023	40.734	-18.928	119	MET O	-10.888	39.838	-18.759
	119	MET CO	-9.810	42.461	-17.055	119	MET CG	-9.880	43.883	-16.582
30	119	MET CD	-8.788	44.963	-17.526	119	MET CD	-9.982	46.061	-18.263
	120	ASP M	-8.904	40.437	-19.584	120	ASP CA	-8.480	39.110	-20.830
	120	ASP C	-7.822	34.390	-18.856	120	ASP O	-8.038	37.109	-18.690
	120	ASP CO	-7.555	39.156	-21.236	120	ASP CG	-8.237	39.730	-22.456
	120	ASP CD1	-7.801	40.706	-23.084	120	ASP CD2	-9.327	39.135	-22.739
	121	VAL M	-7.021	39.117	-18.135	121	VAL CA	-6.224	38.601	-16.974
	121	VAL C	-6.296	39.534	-15.786	121	VAL O	-6.204	40.788	-15.909
	121	VAL CO	-6.755	38.587	-17.496	121	VAL CG1	-3.758	38.176	-16.627
35	121	VAL CG2	-4.787	37.916	-18.046	122	ILE M	-6.318	38.978	-16.598
	122	ILE CA	-6.248	39.799	-13.397	122	ILE C	-5.828	39.262	-12.627
	122	ILE O	-6.829	38.812	-12.469	122	ILE CO	-7.476	39.694	-12.686
	122	ILE CG1	-8.686	40.392	-13.063	122	ILE CG2	-7.221	39.883	-10.954
	122	ILE CD1	-9.976	39.788	-12.393	123	ASN M	-4.263	40.222	-12.110
	123	ASN CA	-3.145	39.854	-11.232	123	ASN C	-3.502	40.404	-9.861
	123	ASN O	-3.708	41.631	-9.833	123	ASN CO	-1.828	40.478	-11.697
40	123	ASN CG	-0.692	40.848	-18.777	123	ASN CD1	-0.863	38.990	-11.018
	123	ASN CD2	-0.346	40.747	-9.728	124	MET M	-3.458	39.604	-8.837
	124	MET CA	-3.658	39.973	-7.438	124	MET C	-2.423	39.683	-6.614

45

50

55

5	124	NET D	-2.306	39.908	-6.893	124	NET C8	-6.943	39.987	-6.893
	124	NET CG	-6.198	40.982	-7.673	124	NET SC	-7.985	39.672	-8.150
	124	NET C1	-7.948	39.993	-7.942	125	SEP M	-1.494	40.696	-6.902
	125	SEP CA	-0.193	40.187	-3.769	125	SEP C	-0.622	40.712	-6.126
	125	SEP D	0.239	41.617	-3.803	125	SEP C8	1.021	41.027	-6.328
	125	SEP CG	1.444	40.496	-7.575	126	LEV M	-1.439	40.076	-3.779
	126	LEV CA	-1.642	40.347	-2.386	126	LEV C	-2.438	39.956	-1.007
	126	LEV D	-2.044	38.136	-2.529	126	LEV C8	-2.791	41.968	-2.410
	126	LEV CG	-3.988	41.447	-3.333	126	LEV C81	-3.278	41.131	-2.578
	126	LEV C82	-4.179	42.760	-4.073	127	GLY M	-2.622	39.082	-0.481
	127	GLY CA	-3.035	37.071	0.193	127	GLY C	-3.176	38.180	2.482
	127	GLY C	-2.446	39.030	2.220	128	GLY M	-4.121	37.443	2.222
	128	GLY CA	-6.679	37.496	2.842	128	GLY C	-6.644	38.036	4.104
	128	GLY D	-6.093	36.198	3.276	129	PRO M	-6.919	39.897	6.492
10	129	PRO CA	-6.671	34.923	0.998	129	PRO C	-6.116	34.886	6.082
	129	PRO D	-6.338	32.187	6.303	129	PRO C8	-6.060	34.684	7.384
	129	PRO CG	-6.419	36.116	7.727	129	PRO C8	-6.239	36.070	6.618
	130	SEP M	-7.081	35.015	6.912	130	SEP CA	-0.470	34.611	6.023
	130	SEP C	-0.218	34.884	4.724	130	SEP C	-0.949	35.881	4.029
	130	SEP C8	-0.849	35.351	7.216	130	SEP C8	-0.723	34.624	6.403
	131	GLY M	-10.083	33.967	4.349	131	GLY CA	-10.624	34.229	3.074
	131	GLY C	-12.205	34.713	3.942	131	GLY D	-12.495	34.722	4.791
15	132	SEP M	-13.940	33.038	2.594	132	SEP CA	-16.407	35.433	3.011
	132	SEP C	-13.289	34.003	1.936	132	SEP D	-16.799	34.986	0.024
	132	SEP CG	-16.890	36.927	3.143	132	SEP CG	-16.693	37.939	3.075
	133	ALA M	-16.847	34.988	2.204	133	ALA CA	-17.807	34.887	1.324
	133	ALA C	-17.410	36.963	0.997	133	ALA D	-17.743	34.437	-1.014
	133	ALA C8	-18.866	33.928	1.996	134	ALA M	-17.683	36.188	0.294
	134	ALA CA	-17.872	37.299	-0.792	134	ALA C	-18.635	37.269	-1.674
	134	ALA D	-16.781	37.985	-2.969	134	ALA C8	-18.263	38.600	-0.187
20	135	LEV M	-16.478	37.229	-1.046	135	LEV CA	-16.197	37.264	-1.804
	135	LEV C	-16.158	36.005	-2.703	135	LEV D	-13.794	36.020	-3.890
	135	LEV CG	-13.038	37.328	-0.798	135	LEV CG	-11.693	37.130	-1.989
	135	LEV C81	-11.440	36.415	-2.292	135	LEV C82	-10.982	36.807	-0.510
	136	LVS M	-16.809	36.823	-2.173	136	LVS CA	-16.943	33.997	-3.013
	136	LVS C	-13.844	33.739	-4.150	136	LVS C	-19.279	33.431	-9.385
	136	LVS CG	-16.993	32.341	-2.186	136	LVS CG	-16.743	31.867	-3.043
	136	LVS C8	-19.083	29.892	-2.134	136	LVS C8	-19.743	28.707	-2.778
25	136	LVS C82	-19.308	28.411	-4.160	137	ALA M	-16.744	34.260	-3.047
	137	ALA CA	-17.795	34.416	-6.893	137	ALA C	-17.338	38.303	-6.045
	137	ALA D	-17.705	35.049	-7.208	137	ALA C8	-19.094	34.941	-4.263
	138	ALA M	-16.529	36.301	-5.729	138	ALA C8	-16.001	37.311	-6.685
	138	ALA C	-16.903	36.696	-7.937	139	ALA D	-16.985	36.043	-0.762
	138	ALA C8	-19.522	38.967	-5.934	139	VAL M	-13.930	35.959	-7.027
	139	VAL CA	-12.946	35.291	-7.037	139	VAL C	-13.623	34.228	-0.720
	139	VAL D	-13.208	34.070	-9.877	139	VAL C8	-11.830	34.671	-6.960
30	139	VAL C81	-10.919	33.056	-7.066	139	VAL C82	-11.078	33.780	-6.293
	140	ASP M	-16.593	33.836	-8.122	140	ASP CA	-19.274	32.494	-0.929
	140	ASP C	-16.823	33.131	-10.084	140	ASP D	-16.080	32.579	-11.190
	140	ASP CG	-16.149	31.949	-8.195	140	ASP CG	-19.388	38.040	-7.184
	140	ASP C81	-16.178	30.603	-7.282	140	ASP C82	-16.139	30.132	-6.329
	141	LVS M	-16.658	34.263	-9.020	141	LVS CA	-17.373	31.006	-10.060
	141	LVS C	-16.373	33.415	-11.946	141	LVS D	-16.780	35.248	-13.111
	141	LVS CG	-18.039	36.275	-10.323	141	LVS CG	-18.084	37.034	-11.306
35	141	LVS C8	-19.686	38.187	-10.536	141	LVS C8	-20.972	39.051	-11.230
	141	LVS C82	-21.138	40.037	-10.273	142	ALA M	-19.147	35.048	-11.866
	142	ALA CA	-16.173	36.192	-12.614	142	ALA C	-19.018	35.010	-13.071
	142	ALA D	-19.770	35.169	-14.789	142	ALA C8	-12.970	36.697	-11.948
	143	VAL M	-19.982	33.886	-12.832	143	VAL CA	-19.160	32.789	-13.650
	143	VAL C	-16.346	32.273	-16.496	143	VAL D	-16.140	31.886	-19.639
	143	VAL CG	-12.831	31.673	-12.714	143	VAL C81	-12.300	30.370	-13.661
40	143	VAL C82	-11.393	32.193	-12.014	144	ALA M	-19.931	32.238	-13.075
	144	ALA CA	-16.744	31.834	-14.643	144	ALA C	-16.928	32.681	-19.961

45

50

55

	144	ALA C	-17.300	32.263	-16.959	144	ALA C8	-17.962	31.968	-13.700
	145	SEP M	-16.507	33.948	-15.701	145	SEP C4	-16.682	34.027	-12.704
	146	SEP C	-15.609	34.773	-17.029	146	SEP D	-15.918	33.321	-18.093
	147	SEP C8	-17.016	36.376	-16.614	147	SEP C8	-16.887	36.993	-19.049
	148	GLY M	-14.577	33.096	-17.965	148	GLY CA	-13.619	33.709	-18.678
5	149	GLY C	-22.273	34.691	-18.385	149	GLY D	-11.620	34.286	-19.266
	150	VAL M	-17.190	35.162	-17.254	150	VAL CA	-10.874	35.856	-16.912
	151	VAL C	-9.850	36.836	-16.223	151	VAL D	-10.171	35.091	-15.486
	152	VAL C8	-11.152	36.977	-15.819	152	VAL C81	-9.866	37.003	-15.578
	153	VAL C82	-12.360	37.913	-16.230	153	VAL M	-8.583	35.818	-16.683
	154	VAL CA	-7.682	34.230	-16.008	154	VAL C	-7.157	34.907	-14.791
	155	VAL D	-6.840	36.133	-14.750	155	VAL C8	-6.273	34.126	-16.958
	156	VAL C81	-5.079	33.483	-16.281	156	VAL C82	-6.990	33.432	-18.262
10	157	VAL M	-7.258	34.355	-13.931	157	VAL CA	-6.987	34.965	-12.249
	158	VAL C	-8.700	34.385	-11.813	158	VAL D	-5.624	33.173	-11.439
	159	VAL C8	-8.224	34.890	-11.315	159	VAL C81	-7.893	35.619	-19.009
	160	VAL C82	-9.486	35.386	-12.094	160	VAL M	-6.722	35.201	-11.494
	161	VAL CA	-3.393	34.987	-10.903	161	VAL C	-3.157	35.623	-9.959
	162	VAL D	-3.892	36.778	-9.400	162	VAL C8	-2.274	35.383	-11.931
	163	VAL C81	-0.973	34.433	-11.461	163	VAL C82	-2.679	34.843	-13.301
	164	ALA M	-2.568	34.946	-9.193	164	ALA CA	-3.361	35.382	-7.287
15	165	ALA C	-1.080	35.036	-8.657	165	ALA D	-0.618	33.889	-6.984
	166	ALA C8	-3.557	35.390	-8.107	166	ALA M	-0.490	35.997	-5.822
	167	ALA CA	0.714	35.638	-9.112	167	ALA C	0.384	34.320	-6.188
	168	ALA D	-0.728	34.464	-3.467	168	ALA C8	1.266	36.697	-4.294
	169	ALA M	1.125	33.302	-3.912	169	ALA CA	0.840	32.258	-2.943
	170	ALA C	0.931	32.725	-1.911	170	ALA D	0.317	32.192	-0.999
	171	ALA C8	1.750	31.038	-3.193	171	GLY M	1.827	33.693	-1.244
	172	GLY CA	2.043	34.211	0.125	172	GLY C	1.519	34.669	0.980
20	173	GLY D	4.189	33.267	-0.118	173	GLY M	3.958	34.788	1.568
	174	ASN CA	0.344	34.787	2.037	174	ASN C	0.399	34.258	3.662
	175	ASN D	6.101	34.829	4.293	175	ASN C8	6.008	36.198	2.984
	176	ASN C8	0.890	36.702	0.900	176	ASN C81	6.123	36.863	-0.934
	177	ASN C82	0.454	37.963	0.382	177	GLU M	4.711	33.168	3.675
	178	GLU CA	4.633	32.837	4.970	178	GLU C	0.322	31.328	0.183
	179	GLU D	0.374	30.637	4.222	179	GLU C8	3.263	31.080	0.180
	180	GLU C8	2.491	32.442	0.368	180	GLU C81	2.394	33.931	6.270
25	181	GLU C82	1.744	34.322	9.317	181	GLU C82	3.186	34.656	7.146
	182	GLY M	6.389	31.057	4.227	182	GLY CA	7.306	30.917	4.387
	183	GLY C	6.603	28.622	4.533	183	GLY D	0.418	28.348	4.089
	184	THR M	7.147	27.793	5.382	184	THR C82	0.070	29.396	3.830
	185	THR C81	0.707	23.487	6.217	185	THR C8	7.864	25.346	5.296
	186	THR CA	6.952	26.487	8.702	186	THR C	6.190	26.480	7.187
	187	THR D	6.470	27.335	7.977	187	THR M	5.338	25.441	7.497
	188	SEP C8	3.141	23.904	10.529	188	SEP C8	3.673	26.185	9.212
30	189	SEP CA	4.033	25.210	0.955	189	SEP C	4.494	23.720	8.944
	190	SEP D	3.339	23.281	0.030	190	GLY M	9.574	22.067	8.835
	191	GLY CA	5.434	21.804	0.095	191	GLY C	4.576	21.049	7.738
	192	GLY D	4.808	21.376	0.555	192	SEP M	3.925	20.318	6.116
	193	SEP CA	2.634	19.777	7.054	193	SEP C	1.477	20.788	4.784
	194	SEP D	0.696	20.347	9.069	194	SEP C8	2.344	18.293	7.271
	195	SEP C8	1.894	18.028	8.559	195	SEP M	1.303	21.841	7.699
	196	SEP C81	0.167	22.725	7.113	196	SEP C	0.430	23.552	9.048
35	197	SEP C	1.533	23.040	9.994	197	SEP C8	-0.213	23.666	8.342
	198	SEP C8	0.104	23.091	9.680	198	SEP M	-0.679	23.921	8.197
	199	SEP CA	-0.611	24.750	3.990	199	SEP C	-0.441	26.177	4.519
	200	SEP D	-1.078	24.348	5.904	200	SEP C8	-1.890	24.642	3.211
	201	SEP C81	-1.982	23.718	2.331	201	THR M	0.387	26.932	3.932
	202	THR CA	0.609	20.340	4.312	202	THR C	0.183	29.286	3.194
	203	THR D	5.483	20.902	3.278	203	THR C8	2.095	28.518	4.818
	204	THR C8	2.984	20.292	3.692	204	THR C82	2.397	27.610	6.001
40	205	VAL M	-0.513	20.742	2.190	205	VAL CA	-0.959	24.342	1.810
	206	VAL C	-2.028	20.943	1.697	206	VAL D	-2.929	20.137	2.280

45

50

55

5	165	VAL C0	-1.319	20.624	-0.161	165	VAL C21	-1.047	20.397	-1.174
	165	VAL C02	-3.210	27.716	-0.091	166	GLY M	-1.910	31.021	1.129
	166	GLY CA	-2.943	32.778	1.676	166	GLY C	-4.098	32.030	0.617
	166	GLY D	-6.124	32.106	-0.396	167	TYR M	-3.034	33.730	0.970
	167	TYR CA	-6.223	34.066	0.113	167	TYR C	-3.993	33.300	-0.006
	167	TYR D	-8.474	36.283	0.004	167	TYR CD	-7.464	34.232	0.964
	167	TYR CG	-7.791	32.964	1.709	167	TYR CD1	-7.208	32.793	1.047
	167	TYR CD2	-8.710	32.116	1.133	167	TYR CF1	-7.567	31.520	3.618
	167	TYR CF2	-9.068	30.933	1.009	167	TYR CF	-8.486	30.671	1.746
	167	TYR DM	-8.082	29.481	3.650	168	PRO M	-6.380	33.499	-1.030
	168	PRO CG	-6.943	36.376	-3.939	168	PRO CD	-6.273	34.762	-1.624
	168	PRO CD	-7.064	35.344	-3.903	168	PRO CA	-7.134	34.437	-2.560
	168	PRO C	-6.398	33.336	-3.170	168	PRO D	-7.097	32.520	-3.913
10	169	GLY M	-3.006	33.193	-3.309	169	GLY CA	-6.446	32.077	-3.027
	169	GLY C	-6.937	30.702	-3.470	169	GLY D	-6.000	29.733	-4.249
	170	LYS M	-3.602	30.379	-2.293	170	LYS CA	-3.036	29.168	-3.113
	170	LYS C	-7.033	28.773	-2.516	170	LYS D	-7.308	27.854	-2.824
	170	LYS CD	-6.246	29.294	-0.384	170	LYS CG	-9.798	28.106	0.933
	170	LYS CD	-6.250	28.289	2.031	170	LYS CF	-5.731	27.271	3.029
	170	LYS M2	-6.239	27.663	3.215	171	TYR M	-7.038	29.616	-3.168
15	171	TYR CA	-3.012	29.043	-3.939	171	TYR C	-8.083	28.300	-3.113
	171	TYR D	-7.760	28.714	-5.928	171	TYR CD	-9.962	30.224	-4.242
	171	TYR CG	-10.497	30.064	-3.047	171	TYR CD1	-11.040	30.303	-1.982
	171	TYR CD2	-10.496	32.374	-3.026	171	TYR CF1	-11.920	31.003	-0.067
	171	TYR CF2	-10.941	31.008	-1.936	171	TYR CF	-11.920	32.396	-0.006
	171	TYR DM	-12.008	33.119	0.170	172	PRO M	-9.297	27.894	-3.374
	172	PRO CA	-9.093	26.617	-6.396	172	PRO C	-9.233	27.166	-7.000
	172	PRO D	-8.323	26.784	-0.881	172	PRO CD	-10.167	29.329	-6.313
20	172	PRO CG	-10.650	29.271	-9.096	172	PRO CD	-10.364	26.669	-4.816
	173	SER M	-10.017	28.167	-0.019	173	SER CA	-10.220	28.818	-0.330
	173	SER C	-9.025	29.773	-9.393	173	SER D	-8.966	30.233	-10.742
	173	SER CD	-11.928	29.023	-9.481	173	SER CG	-11.393	30.946	-8.406
	174	VAL M	-8.162	29.944	-8.614	174	VAL CA	-7.093	30.091	-8.093
	174	VAL C	-9.794	30.131	-9.068	174	VAL D	-5.812	29.192	-8.344
	174	VAL CD	-6.891	31.773	-7.396	174	VAL CG1	-5.796	32.037	-7.617
	174	VAL CG2	-8.220	32.503	-7.323	175	ILE M	-6.911	30.729	-9.001
25	175	ILE CA	-3.869	30.156	-10.024	175	ILE C	-2.714	30.736	-0.094
	175	ILE D	-2.450	31.098	-8.953	175	ILE CD	-2.993	30.524	-11.619
	175	ILE CG1	-3.857	29.978	-12.924	175	ILE CG2	-1.491	30.089	-11.512
	175	ILE CF1	-3.692	30.929	-13.946	176	ALA M	-2.220	30.020	-7.929
	176	ALA CA	-1.335	30.917	-6.870	176	ALA C	0.120	30.301	-7.310
	176	ALA D	0.433	29.215	-7.038	176	ALA CD	-1.639	29.830	-8.541
	177	VAL M	0.064	31.410	-7.180	177	VAL CA	2.261	31.834	-7.096
	177	VAL C	3.223	31.693	-6.473	177	VAL D	3.176	32.057	-9.721
30	177	VAL CD	2.439	32.607	-8.793	177	VAL CG1	3.042	32.067	-9.392
	177	VAL CG2	1.374	32.332	-9.843	178	GLY M	6.077	30.634	-6.398
	178	GLY CA	5.100	30.703	-3.339	178	GLY C	6.446	31.233	-6.074
	178	GLY D	6.491	31.435	-7.206	179	ALA M	7.012	31.467	-9.287
	179	ALA CA	0.715	32.037	-3.839	179	ALA C	9.939	31.090	-8.779
	179	ALA C	10.198	30.481	-4.719	179	ALA CD	9.025	33.251	-6.973
	180	VAL M	10.639	31.162	-6.085	180	VAL CA	11.970	30.472	-6.991
35	180	VAL C	13.948	31.805	-7.171	180	VAL D	12.712	32.691	-7.427
	180	VAL CD	12.073	29.914	-8.166	180	VAL CG1	11.271	28.231	-7.833
	180	VAL CG2	11.673	30.129	-9.900	181	ASP M	14.267	31.209	-6.000
	181	ASP CA	13.431	32.100	-7.039	181	ASP C	10.942	31.804	-8.462
	181	ASP D	19.339	31.090	-9.292	181	ASP CD	16.446	31.921	-8.916
	181	ASP CG	17.120	30.934	-9.971	181	ASP CD1	17.103	29.793	-6.972
	181	ASP CD2	17.000	30.296	-4.007	182	SER M	17.087	32.306	-8.047
	182	SER CA	17.622	32.214	-10.191	182	SER C	10.193	30.817	-10.494
40	182	SER D	10.363	30.492	-11.070	182	SER CD	10.678	33.313	-10.466
	182	SER CG	10.016	34.961	-10.675	183	SER M	10.238	30.042	-9.423
	183	SER CA	10.716	28.043	-9.444	183	SER C	17.991	27.614	-9.947
	183	SER D	17.039	26.415	-9.397	183	SER CD	10.256	28.923	-8.007

45

50

55

EP 0 251 446 B1

	103	SEB	OC	29.589	28.613	-0.251	104	ASB	N	16.373	28.004	-0.002
	104	ASB	CA	15.144	27.317	-0.950	104	ASB	C	16.931	26.720	-0.197
	104	ASB	O	14.138	25.759	-0.097	104	ASB	CB	15.014	26.341	-10.722
	104	ASB	CC	14.993	26.998	-12.076	104	ASB	CC1	14.798	26.104	-12.277
5	104	ASB	CC2	14.952	24.710	-13.076	103	GLN	N	13.942	27.247	-7.199
	103	GLN	CA	13.276	26.646	-0.033	103	GLN	C	14.280	27.694	-5.103
	103	GLN	O	14.199	25.726	-0.396	103	GLN	CB	16.599	26.568	-9.101
	103	GLN	CC	16.539	26.242	-3.614	103	GLN	CC	18.911	26.182	-9.206
	103	GLN	CC1	18.864	25.799	-6.061	103	GLN	CC2	18.266	26.386	-1.934
	106	ABC	N	13.278	26.938	-4.448	106	ABC	CA	12.183	27.774	-3.041
	106	ABC	C	12.780	28.782	-2.866	106	ABC	O	13.698	28.384	-3.093
	106	ABC	CB	11.315	26.043	-3.116	106	ABC	CC	18.214	27.471	-2.161
10	106	ABC	CC	9.487	26.337	-1.468	106	ABC	CC1	9.866	26.333	-0.117
	106	ABC	CC2	9.961	26.879	1.059	106	ABC	CC3	9.367	27.050	1.058
	106	ABC	CC4	10.966	26.271	1.703	107	ALA	N	12.294	28.009	-2.053
	107	ALA	CA	12.728	31.064	-1.095	107	ALA	C	12.262	28.604	-0.517
	107	ALA	O	11.193	30.043	-0.387	107	ALA	CB	12.144	27.692	-2.544
	108	SEB	N	13.931	30.770	0.549	108	SEB	CA	12.671	30.286	1.868
	108	SEB	C	11.356	30.847	2.412	108	SEB	O	18.740	30.111	3.212
	108	SEB	CB	13.767	30.456	2.932	108	SEB	CC	14.137	31.024	2.841
15	109	PME	N	10.943	32.010	1.974	109	PME	CA	9.697	32.688	2.418
	109	PME	C	8.499	32.198	1.609	109	PME	O	7.989	32.856	2.011
	109	PME	CB	9.787	34.217	2.243	109	PME	CC	10.117	34.098	0.867
	109	PME	CC1	9.147	34.030	-0.121	109	PME	CC2	11.418	35.114	0.967
	109	PME	CC2	9.483	33.187	-1.411	109	PME	CC3	11.769	35.949	-0.701
	109	PME	CC3	10.786	35.066	-1.728	109	PME	CC4	9.703	31.926	0.499
	100	SEB	CA	7.626	31.094	-0.391	100	SEB	C	6.663	30.162	0.328
	100	SEB	O	7.034	29.083	0.066	100	SEB	CB	8.181	30.990	-1.708
20	100	SEB	CC	7.136	30.337	-2.618	101	SEB	N	8.388	30.951	0.326
	101	SEB	CA	4.341	29.696	0.987	101	SEB	C	4.261	28.330	0.223
	101	SEB	O	4.543	28.268	-0.093	101	SEB	CB	3.915	30.611	0.911
	101	SEB	CC	2.729	31.283	1.934	102	VAL	N	3.786	27.918	0.920
	102	VAL	CA	3.629	23.932	0.391	102	VAL	C	2.284	25.291	0.686
	102	VAL	O	1.899	25.698	1.998	102	VAL	CB	4.781	29.127	1.088
	102	VAL	CC1	6.144	29.727	0.722	102	VAL	CC2	4.617	25.104	2.592
	103	GLY	N	1.938	24.172	0.047	103	GLY	CA	8.629	29.944	0.416
25	103	GLY	C	0.881	23.029	-0.901	103	GLY	O	8.530	23.244	-2.015
	104	PRC	N	-1.023	22.389	-0.722	104	PRC	CA	-1.662	21.691	-1.073
	104	PRC	C	-1.237	22.603	-2.914	104	PRC	O	-2.603	22.244	-6.085
	104	PRC	CB	-2.769	20.783	-1.210	104	PRC	CC	-2.311	20.622	0.213
	104	PRC	CC1	-1.633	21.954	0.578	105	GLU	N	-2.822	23.793	-2.439
	105	GLU	CA	-3.165	24.080	-3.232	105	GLU	C	-2.095	25.631	-4.058
	105	GLU	O	-2.516	26.398	-4.936	105	GLU	CB	-4.043	26.786	-3.478
	105	GLU	CC	-4.942	25.134	-1.435	105	GLU	CC	-4.313	24.860	-0.188
30	105	GLU	CC1	-3.120	24.960	0.163	105	GLU	CC2	-5.138	24.520	0.785
	106	LEU	N	-0.829	23.264	-3.070	106	LEU	CA	8.241	25.929	-4.664
	106	LEU	C	0.228	25.376	-6.059	106	LEU	O	8.305	24.121	-6.193
	106	LEU	CB	1.340	25.739	-3.084	106	LEU	CC	3.770	26.178	-4.643
	106	LEU	CC1	2.739	27.716	-4.439	106	LEU	CC2	4.827	25.721	-3.911
	107	ASP	N	0.140	26.208	-7.093	107	ASP	CA	8.032	25.774	-8.480
	107	ASP	C	1.387	25.738	-9.293	107	ASP	O	1.033	24.734	-9.914
	107	ASP	CB	-1.067	26.998	-9.191	107	ASP	CC	-2.686	26.351	-8.949
35	107	ASP	CC1	-1.864	25.155	-8.394	107	ASP	CC2	-3.033	27.327	-8.088
	108	VAL	N	2.013	26.089	-0.344	108	VAL	CA	3.284	26.970	-10.209
	108	VAL	C	4.157	27.030	-9.314	108	VAL	O	3.752	28.099	-8.587
	108	VAL	CB	2.894	27.476	-11.637	108	VAL	CC1	1.938	26.726	-12.937
	108	VAL	CC2	2.337	28.019	-11.484	109	MEY	N	9.374	27.016	-10.016
	109	MEY	CA	6.439	28.802	-9.498	109	MEY	C	6.849	29.010	-10.578
	109	MEY	O	6.696	29.518	-11.793	109	MEY	CB	7.660	27.970	-9.877
	109	MEY	CC	7.363	26.849	-8.139	109	MEY	CC	6.753	27.649	-8.568
40	109	MEY	CC1	8.227	27.753	-8.987	200	ALA	N	7.626	30.942	-10.103
	200	ALA	CA	7.991	31.929	-11.033	200	ALA	C	9.888	32.068	-10.272
	200	ALA	O	8.327	32.674	-9.060	200	ALA	CB	8.932	32.878	-11.638

45

50

55

	201	PAC M	9.927	32.499	-10.991	201	PAC CA	11.813	34.130	-10.230
	201	PAC C	10.430	35.127	-9.230	201	PAC D	9.979	35.987	-9.681
	201	PAC C8	11.017	36.723	-11.490	201	PAC CG	11.392	36.040	-12.670
	201	PAC CD	9.941	33.616	-12.405	202	GLY M	10.923	31.784	-8.021
	202	GLY CA	10.473	36.204	-7.844	202	GLY C	11.780	36.636	-6.135
	202	GLY D	11.332	37.124	-4.979	203	VAL M	12.813	36.393	-6.613
	203	VAL CA	13.040	36.929	-9.716	203	VAL C	14.706	38.017	-6.669
	203	VAL C	13.133	37.731	-7.593	203	VAL CG	14.814	38.680	-5.351
	203	VAL CG1	16.096	36.106	-4.612	203	VAL CG2	14.879	34.741	-4.376
	204	SER M	14.061	39.182	-5.839	204	SER CA	15.572	40.281	-6.487
	204	SER C	15.067	40.619	-7.872	204	SER C	15.786	40.683	-8.889
	204	SER CG	17.037	39.976	-6.326	204	SER CG	17.732	41.186	-6.672
	205	ILE M	13.771	40.963	-8.000	205	ILE CA	13.969	41.234	-9.228
	205	ILE C	13.207	42.749	-8.478	205	ILE D	12.678	43.490	-8.640
	205	ILE CG	11.532	40.833	-9.144	205	ILE CG1	11.436	39.336	-8.810
	205	ILE CG2	10.899	41.281	-10.467	205	ILE CG1	12.237	38.412	-9.771
	206	GLN M	13.936	41.939	-10.689	206	GLN CA	14.204	44.317	-10.834
	206	GLN C	13.902	44.978	-11.630	206	GLN D	12.669	44.318	-12.621
	206	GLN CG	15.435	44.708	-11.740	206	GLN CG	16.684	44.163	-10.980
	206	GLN CD	17.285	45.145	-10.007	206	GLN CD1	16.320	44.936	-9.353
	206	GLN CD2	16.536	46.260	-9.857	207	SER M	12.359	46.064	-11.216
	207	SER CA	11.217	46.571	-11.987	207	SER C	11.089	40.093	-11.740
	207	SER D	11.919	48.657	-11.004	207	SER CG	9.918	49.053	-11.569
	207	SER CG	8.993	46.036	-12.613	208	THR M	10.854	48.064	-12.326
	208	THR CG2	9.171	50.339	-14.734	208	THR CG1	7.570	49.414	-13.144
	208	THR CG	8.620	50.415	-13.357	208	THR CA	9.678	50.092	-12.179
	208	THR C	9.197	50.489	-10.803	208	THR D	8.423	49.807	-10.049
	209	LEU M	9.636	51.613	-10.228	209	LEU CA	9.192	52.158	-8.930
	209	LEU C	8.473	53.610	-8.262	209	LEU D	9.140	54.227	-10.222
	209	LEU CG	10.335	52.192	-7.958	209	LEU CG	10.804	59.816	-7.416
	209	LEU CD1	11.968	51.114	-6.472	209	LEU CD2	9.607	50.282	-8.649
	210	PRO M	7.790	54.139	-8.444	210	PRO CA	7.273	55.517	-8.649
	210	PRO C	8.383	56.573	-8.639	210	PRO D	9.491	56.645	-8.104
	210	PRO CG	6.302	55.733	-7.917	210	PRO CG	6.004	54.379	-6.944
	210	PRO CD	7.193	53.491	-7.271	211	GLY M	8.077	57.665	-8.339
	211	GLY CA	9.049	58.763	-9.410	211	GLY C	10.894	58.454	-10.490
	211	GLY D	11.176	59.005	-10.289	212	ASN M	9.891	57.770	-11.057
	212	ASN CA	10.903	57.422	-12.643	212	ASN C	12.039	58.753	-12.056
	212	ASN C	13.188	57.361	-12.620	212	ASN CG	11.224	58.393	-13.499
	212	ASN CG	11.803	58.189	-14.814	212	ASN CD1	11.053	57.054	-15.323
	212	ASN CD2	12.273	59.159	-15.376	213	LYS M	11.803	59.749	-11.247
	213	LYS CA	12.810	54.946	-10.337	213	LYS C	12.668	53.639	-10.066
	213	LYS D	11.779	53.039	-11.613	213	LYS CG	12.769	55.241	-9.838
	213	LYS CG	13.204	56.694	-8.767	213	LYS CD	13.246	57.030	-7.312
	213	LYS CE	14.159	58.218	-6.870	213	LYS ME	15.040	58.705	-7.921
	214	TYR M	13.631	52.703	-10.444	214	TYR CA	13.803	51.246	-10.722
	214	TYR C	16.383	50.600	-9.689	214	TYR D	19.211	51.293	-8.817
	214	TYR CG	14.641	50.981	-11.984	214	TYR CG	14.130	51.621	-13.746
	214	TYR CD1	14.689	52.047	-13.678	214	TYR CD2	13.129	51.063	-14.014
	214	TYR CE1	14.230	53.479	-14.814	214	TYR CD2	12.654	51.669	-15.178
	214	TYR CE2	13.204	52.095	-15.850	214	TYR DM	12.756	53.438	-16.496
	215	GLY M	14.838	49.847	-9.158	215	GLY CA	14.622	46.772	-7.905
	215	GLY C	14.130	47.328	-7.749	215	GLY D	13.249	46.917	-8.321
	216	ALA M	14.810	46.839	-8.831	216	ALA CA	14.454	48.209	-8.781
	216	ALA C	13.832	44.922	-8.712	216	ALA D	13.948	45.327	-6.479
	216	ALA CG	19.715	44.354	-8.897	217	TYR M	12.788	43.982	-5.975
	217	TYR CA	11.964	43.688	-4.640	217	TYR C	12.033	41.920	-4.547
	217	TYR D	12.202	41.642	-9.636	217	TYR CG	10.473	49.862	-4.570
	217	TYR CG	10.117	49.291	-4.214	217	TYR CD1	10.846	49.991	-3.236
	217	TYR CD2	9.016	49.933	-4.789	217	TYR CE1	10.459	47.267	-2.790
	217	TYR CE2	8.634	47.219	-4.381	217	TYR CE2	9.358	47.882	-3.391
	217	TYR DM	8.933	49.160	-2.989	218	ASN M	11.780	41.386	-3.591
	218	ASN CA	11.640	39.942	-3.227	218	ASN C	10.204	39.636	-2.769

EP 0 251 446 B1

	218	ASL D	0.763	40.347	-1.017	218	ASL CD	12.053	30.340	-2.154
	218	ASL CG	14.031	30.866	-2.343	218	ASL DD1	14.612	30.700	-3.421
	218	ASL MD2	14.660	30.644	-3.105	219	GLY H	0.670	30.934	-2.209
	219	GLY CA	0.302	30.132	-2.669	219	GLY C	7.570	37.304	-3.601
5	219	GLY D	7.073	37.002	-4.076	220	YMR H	6.561	26.630	-3.203
	220	YMR CA	0.697	35.936	-4.176	220	YMR C	4.870	37.044	-4.064
	220	YMR D	4.417	36.742	-5.930	220	YMR CD	4.025	34.010	-3.026
	220	YMR DD1	4.136	30.543	-2.451	220	YMR CD2	5.704	33.094	-2.980
	221	SER H	4.738	30.230	-4.353	221	SER CA	3.004	39.201	-5.160
	221	SER C	4.760	39.641	-6.303	221	SER D	4.117	40.208	-7.277
	221	SER CD	3.329	40.303	-4.546	221	SER CD	3.435	40.202	-3.140
	222	MEI H	0.040	39.309	-6.605	222	MEI CE	6.471	42.771	-9.173
10	222	MEI SD	7.760	41.333	-4.993	222	MEI CG	0.506	41.399	-6.602
	222	MEI CD	0.351	40.010	-7.218	222	MEI CA	6.010	39.670	-7.630
	222	MEI C	6.077	30.435	-0.567	222	MEI D	7.004	30.567	-9.775
	223	ALA H	6.054	37.244	-0.041	223	ALA CA	6.469	26.020	-0.005
	223	ALA C	5.200	34.068	-0.707	223	ALA D	3.133	35.940	-10.929
	223	ALA CD	0.509	34.007	-7.023	224	SER H	4.070	34.360	-9.030
	224	SER CA	2.798	36.680	-0.700	224	SER C	2.661	37.161	-11.039
	224	SER D	2.145	36.593	-12.057	224	SER CD	1.001	36.995	-0.603
15	224	SER CD	0.492	36.099	-0.157	225	PRD H	3.150	30.411	-11.159
	225	PRD CA	3.095	39.130	-12.439	225	PRD C	3.764	30.469	-13.074
	225	PRD D	3.406	30.650	-14.004	225	PRD CD	3.693	40.911	-13.054
	225	PRD CG	4.411	40.402	-10.764	225	PRD CD	3.735	39.224	-10.054
	226	MIS H	4.709	37.026	-13.299	226	MIS CA	5.446	34.070	-14.362
	226	MIS C	4.410	35.947	-15.061	226	MIS D	4.425	35.009	-16.293
	226	MIS CD	6.008	34.046	-13.745	226	MIS CG	7.014	36.059	-13.358
	226	MIS DD1	0.040	37.688	-12.170	226	MIS CD2	0.003	37.110	-14.167
20	226	MIS CD1	9.270	30.052	-12.236	226	MIS ME2	0.771	37.066	-13.443
	227	VAL H	3.593	35.366	-14.195	227	VAL CA	2.603	34.300	-14.721
	227	VAL C	1.479	35.197	-15.421	227	VAL D	1.016	34.773	-16.490
	227	VAL CD	2.303	33.044	-13.610	227	VAL CD1	1.076	32.476	-14.246
	227	VAL CD2	3.204	32.665	-12.891	228	ALA H	1.003	36.242	-14.014
	228	ALA CA	0.011	37.109	-15.517	228	ALA C	0.543	37.930	-16.000
	228	ALA D	-0.233	37.433	-17.020	228	ALA CD	-0.307	30.353	-14.660
	229	GLY H	1.701	30.028	-16.941	229	GLY CA	2.352	30.400	-18.239
25	229	GLY C	2.420	37.197	-19.187	229	GLY D	2.109	37.373	-20.304
	230	ALA H	2.711	35.908	-18.446	230	ALA CA	2.794	34.001	-19.346
	230	ALA C	1.424	34.000	-20.133	230	ALA D	1.380	34.203	-21.343
	230	ALA CD	3.200	33.624	-18.700	231	ALA H	0.305	34.623	-19.320
	231	ALA CA	-1.010	34.416	-19.744	231	ALA C	-1.256	35.423	-20.064
	231	ALA D	-1.900	35.050	-21.052	231	ALA CD	-1.032	34.664	-20.049
	232	ALA H	-0.770	36.657	-20.721	232	ALA CA	-1.013	37.003	-21.792
	232	ALA C	-0.201	37.204	-23.070	232	ALA D	-0.041	37.901	-24.187
30	232	ALA CD	-0.742	39.121	-21.377	233	LEU H	0.935	36.724	-22.967
	233	LEU CA	1.617	34.293	-24.209	233	LEU C	0.621	35.169	-24.000
	233	LEU D	0.696	35.231	-26.111	233	LEU CD	3.063	35.077	-23.907
	233	LEU CG	3.996	36.994	-23.453	233	LEU CD1	5.239	36.342	-22.921
	233	LEU CD2	4.241	37.853	-24.600	234	ILE H	0.537	34.199	-24.047
	234	ILE CD1	0.306	30.664	-21.657	234	ILE CG1	0.454	31.223	-23.103
	234	ILE D	-0.011	32.014	-23.570	234	ILE CG2	-1.003	30.900	-24.091
	234	ILE CA	-0.406	33.076	-24.644	234	ILE C	-1.621	33.997	-23.434
35	234	ILE D	-1.003	33.144	-24.144	235	LEU H	-2.390	34.463	-24.770
	235	LEU CA	-3.096	35.020	-25.423	235	LEU C	-3.250	35.043	-26.071
	235	LEU D	-4.109	35.914	-27.000	235	LEU CD	-4.432	35.765	-24.370
	235	LEU CG	-5.160	34.009	-23.342	235	LEU CD1	-5.052	35.003	-22.165
	235	LEU CD2	-6.252	34.130	-24.120	236	SER H	-2.094	36.430	-26.790
	236	SER CA	-1.764	37.237	-27.000	236	SER C	-1.491	36.292	-29.144
	236	SER D	-1.746	36.634	-30.290	236	SER CD	-0.633	38.234	-27.733
	236	SER CD	0.099	37.571	-27.982	237	LYS H	-1.046	35.067	-28.002
40	237	LYS CA	-0.046	34.005	-29.952	237	LYS C	-2.113	33.277	-30.260
	237	LYS D	-2.370	32.951	-31.444	237	LYS CD	0.272	33.112	-29.551
	237	LYS CG	0.677	32.240	-30.716	237	LYS CD	2.020	31.035	-30.442

45

50

55

5	237	LVI CE	2.345	30.762	-31.729	237	LVI M2	3.023	20.048	-31.006
	238	M13 M	-2.951	31.910	-20.312	238	M13 C0	-4.360	32.163	-20.370
	239	M13 C	-5.334	32.899	-20.497	239	M13 D	-5.713	32.984	-27.962
	239	M13 C0	-3.968	30.962	-20.511	239	M13 C6	-3.000	20.021	-20.237
	239	M13 M02	-1.707	29.670	-20.093	239	M13 C02	-3.137	20.390	-30.394
10	239	M13 C01	-1.086	20.031	-20.662	239	M13 M02	-1.940	20.600	-30.199
	239	P00 M	-3.048	33.917	-20.363	239	P00 CA	-6.908	36.770	-20.773
	239	P00 C	-8.204	34.012	-20.332	239	P00 D	-0.949	34.910	-27.662
	239	P00 C0	-7.010	35.977	-20.713	239	P00 C6	-6.666	39.294	-31.027
	239	P00 C0	-9.436	36.430	-20.660	240	ASN M	-3.306	32.969	-20.227
15	240	ASN CA	-9.929	32.041	-20.216	240	ASN C	-9.500	31.100	-27.980
	240	ASN D	-10.940	30.610	-27.976	240	ASN C0	-9.493	31.249	-30.333
	240	ASN C6	-7.971	30.027	-30.009	240	ASN C01	-7.000	31.990	-31.147
	240	ASN M02	-7.670	29.909	-30.976	241	T0P M	-8.391	31.006	-27.304
	241	T0P CA	-8.304	30.124	-26.170	241	T0P C	-9.106	30.630	-24.936
20	241	T0P D	-9.043	31.833	-24.606	241	T0P C0	-6.070	29.030	-25.670
	241	T0P C6	-6.094	28.903	-26.957	241	T0P C01	-6.300	28.433	-27.810
	241	T0P C02	-6.039	28.976	-26.193	241	T0P M01	-3.362	27.947	-20.211
	241	T0P C02	-6.414	27.676	-27.216	241	T0P C02	-4.097	20.606	-24.981
	241	T0P C02	-3.193	26.706	-27.174	241	T0P C10	-2.912	27.667	-24.943
25	241	T0P C02	-2.470	26.073	-26.003	242	TMR M	-0.727	20.701	-24.162
	242	TMR C0	-10.458	30.119	-22.911	242	TMR C	-9.469	30.176	-21.747
	242	TMR D	-8.333	29.674	-21.937	242	TMR C0	-11.070	20.032	-22.673
	242	TMR C01	-10.037	27.706	-22.476	242	TMR C02	-12.404	20.007	-23.095
	243	ASN M	-9.946	30.639	-20.611	243	ASN M02	-11.707	20.004	-21.767
30	243	ASN C01	-11.463	31.810	-10.750	243	ASN C6	-11.093	31.131	-17.983
	243	ASN C0	-9.708	31.830	-10.332	243	ASN CA	-0.053	30.731	-10.664
	243	ASN C	-8.637	29.303	-10.010	243	ASN D	-7.093	29.136	-10.440
	244	TMR M	-9.064	28.162	-10.233	244	TMR CA	-9.381	24.934	-19.050
	244	TMR C	-8.133	26.393	-10.002	244	TMR D	-7.324	25.787	-19.111
35	244	TMR C0	-10.665	26.000	-10.494	244	TMR C01	-11.735	26.670	-10.684
	244	TMR C02	-10.903	26.995	-19.197	245	GLN M	-8.082	26.716	-21.073
	245	GLN CA	-6.964	26.362	-21.962	245	GLN C	-5.647	27.020	-21.520
	245	GLN D	-6.173	26.393	-21.447	245	GLN C0	-7.330	26.999	-23.397
	245	GLN C6	-8.265	29.326	-23.909	245	GLN C0	-8.493	29.073	-23.428
40	245	GLN M01	-9.306	26.769	-20.727	245	GLN M02	-7.765	28.512	-24.370
	246	VAL M	-5.697	28.304	-21.210	246	VAL CA	-6.477	29.040	-20.770
	246	VAL C	-3.936	28.662	-19.467	246	VAL D	-3.709	20.227	-10.361
	246	VAL C0	-4.779	30.959	-20.671	246	VAL C01	-3.944	31.272	-20.027
	246	VAL C02	-5.169	31.130	-21.959	247	ARC M	-4.707	20.240	-10.662
45	247	ARC CA	-4.380	27.714	-17.160	247	ARC C	-3.770	26.292	-17.360
	247	ARC D	-2.709	29.985	-16.764	247	ARC C0	-3.533	27.667	-16.149
	247	ARC C6	-6.987	27.093	-14.092	247	ARC C0	-6.096	27.170	-13.793
	247	ARC M01	-5.440	26.757	-12.946	247	ARC C2	-5.093	26.066	-11.313
	247	ARC M02	-7.064	27.404	-11.210	247	ARC M02	-5.177	26.620	-10.270
50	248	SER M	-4.610	29.903	-10.191	248	SER CA	-6.039	24.131	-10.626
	248	SER C	-2.637	24.096	-10.072	248	SER D	-1.048	23.293	-10.003
	248	SER C0	-3.034	23.409	-19.372	248	SER C0	-6.146	23.090	-10.032
	249	SER M	-2.300	24.093	-20.136	249	SER C0	-1.223	24.074	-20.051
	249	SER C	-0.071	29.302	-19.940	249	SER D	1.026	24.700	-20.049
55	249	SER C0	-1.369	29.750	-22.000	249	SER C0	-0.300	25.610	-22.956
	250	LEU M	-0.209	26.333	-19.160	250	LEU C02	1.024	29.014	-10.222
	250	LEU C01	-0.373	20.433	-17.260	250	LEU C6	0.352	20.430	-10.191
	250	LEU C0	0.170	20.003	-17.903	250	LEU CA	0.710	26.037	-10.216
	250	LEU C	1.092	29.694	-17.265	250	LEU C	2.293	29.421	-17.032
60	251	GLN M	0.048	29.007	-16.714	251	GLN M02	-2.700	29.512	-12.237
	251	GLN C01	-2.019	23.424	-12.930	251	GLN C0	-2.940	24.950	-13.034
	251	GLN C6	-1.210	24.014	-13.994	251	GLN C0	-0.057	23.621	-14.077
	251	GLN CA	0.981	23.941	-19.743	251	GLN C	0.959	22.664	-14.361
	251	GLN D	1.743	22.014	-19.616	252	ASN M	0.633	22.394	-17.990
65	252	ASN CA	1.002	21.204	-19.702	252	ASN C	2.394	21.399	-10.991
	252	ASN D	1.500	20.442	-19.760	252	ASN C0	0.004	20.700	-10.292
	252	ASN C6	-1.036	19.926	-19.973	252	ASN C01	-0.036	19.393	-17.002

28

5	231	AS4	DC2	-2.234	19.934	-19.161	233	Y44	N	9.018	22.501	-10.921
	231	Y44	CA	4.234	22.717	-19.717	233	Y44	C	9.201	23.247	-10.918
	231	Y44	D	0.341	23.733	-19.623	233	Y44	CB	4.006	23.672	-20.992
	233	Y44	CC1	3.393	24.037	-20.428	233	Y44	CC2	3.147	23.136	-22.032
	234	Y44	N	9.218	21.177	-17.851	234	Y44	CA	4.214	23.612	-16.508
	234	Y44	C	7.466	22.700	-16.612	234	Y44	D	7.402	23.980	-17.093
	234	Y44	CB	9.664	23.930	-15.132	234	Y44	CC1	9.129	22.170	-16.040
	234	Y44	CC2	4.330	24.944	-14.802	235	Y44	N	8.499	23.004	-16.476
	235	Y44	CA	9.771	22.994	-15.017	235	Y44	C	9.021	23.201	-16.414
	235	Y44	D	9.439	22.786	-13.474	235	Y44	CB	11.080	23.481	-15.897
10	235	Y44	CC1	11.002	23.709	-17.321	235	Y44	CC2	12.286	22.620	-15.450
	236	LY4	N	9.606	26.702	-14.314	236	LY4	CA	9.364	20.063	-13.010
	236	LY4	C	10.322	20.393	-12.063	236	LY4	D	11.662	20.274	-12.992
	236	LY4	CB	9.024	10.990	-13.249	236	LY4	CC	9.018	17.005	-11.021
	236	LY4	CD	10.286	16.968	-11.777	236	LY4	CE	10.732	19.940	-10.613
	236	LY4	CE	9.243	14.969	-11.054	237	LY4	N	10.212	20.074	-10.874
	237	LY4	CA	11.272	21.031	-9.993	237	LY4	C	11.230	20.232	-9.614
	237	LY4	C	12.096	27.865	-7.732	237	LY4	CB	11.107	22.947	-9.522
	237	LY4	CC	11.297	23.620	-10.868	237	LY4	CD1	11.248	25.003	-9.921
	237	LY4	CD2	12.078	23.468	-11.325	238	GL4	N	10.631	19.282	-6.288
15	238	GL4	CA	16.002	14.793	-6.079	238	GL4	C	9.168	10.703	-6.373
	238	GL4	C	8.221	10.956	-7.202	239	AS4	N	9.024	10.282	-8.190
	239	AS4	CA	7.757	17.896	-4.910	239	AS4	C	6.619	17.941	-4.709
	239	AS4	D	4.059	20.039	-4.214	239	AS4	CB	7.096	17.940	-3.983
	239	AS4	CC	4.701	17.128	-2.243	239	AS4	DD1	5.611	17.927	-2.394
	239	AS4	CC2	7.098	16.299	-1.121	240	SR4	N	5.560	10.610	-5.312
	240	SR4	CA	4.081	10.987	-9.529	240	SR4	C	4.946	20.367	-6.289
	240	SR4	D	2.300	21.513	-4.646	240	SR4	CB	9.345	10.919	-6.289
	240	SR4	CC	2.743	17.937	-3.668	241	P44	N	4.241	19.778	-3.112
	241	P44	CA	3.831	20.468	-1.085	241	P44	C	4.544	21.046	-1.863
20	241	P44	D	3.944	22.040	-1.432	241	P44	CB	4.053	19.749	-1.563
	241	P44	CC	3.849	20.337	0.719	241	P44	CD1	2.204	20.163	1.125
	241	P44	CC2	4.601	21.045	1.538	241	P44	CE	1.797	20.717	2.315
	241	P44	CE2	3.945	21.602	2.748	241	P44	CC	2.005	21.669	3.114
	242	Y44	N	5.776	21.708	-2.305	242	Y44	CA	4.688	22.914	-2.251
	242	Y44	C	6.020	23.689	-3.849	242	Y44	D	7.201	24.053	-3.393
	242	Y44	CB	8.122	22.439	-1.031	242	Y44	CC	8.146	21.892	-0.454
	242	Y44	CD1	8.094	20.434	-0.361	242	Y44	CD2	8.149	22.669	0.498
	242	Y44	CE1	8.062	19.873	0.882	242	Y44	CE2	8.134	22.069	1.962
	242	Y44	CE	8.069	20.672	2.018	242	Y44	DE	7.941	20.079	3.205
25	243	Y44	N	6.626	23.104	-4.693	243	Y44	CA	6.812	23.693	-0.022
	243	Y44	C	8.024	23.680	-4.956	243	Y44	D	9.781	24.111	-0.111
	243	Y44	CB	7.928	22.768	-6.681	243	Y44	CC	9.279	23.035	-0.048
	243	Y44	CC1	10.964	24.044	-6.637	243	Y44	CC2	9.000	22.342	-4.993
	243	Y44	CE1	11.393	24.328	-6.160	243	Y44	CE2	11.062	22.660	-6.491
	243	Y44	CE	11.038	23.610	-5.184	243	Y44	DE	13.063	23.949	-6.907
	244	GL4	N	4.671	23.163	-6.316	244	GL4	CA	3.301	23.064	-7.412
	244	GL4	C	3.047	22.196	-8.936	244	GL4	D	4.647	21.274	-8.369
	244	LY4	N	3.436	22.477	-9.734	244	LY4	CA	3.034	21.798	-10.971
	245	LY4	C	9.188	21.232	-11.464	245	LY4	D	5.094	21.863	-12.306
30	245	LY4	CB	2.739	22.071	-12.044	245	LY4	CC	1.400	21.963	-11.305
	245	LY4	CD	0.710	20.940	-12.079	245	LY4	CE	-0.602	20.494	-11.391
	245	LY4	CE	-3.078	23.737	-12.489	246	GL4	N	5.707	23.126	-10.917
	246	GL4	C	7.120	21.923	-11.923	246	GL4	C	7.193	23.052	-11.010
	246	GL4	D	6.177	23.793	-11.648	247	LY4	N	8.262	25.236	-12.490
	247	LY4	CA	6.490	24.640	-13.097	247	LY4	C	7.004	26.771	-14.437
	247	LY4	D	7.993	25.909	-15.208	247	LY4	CB	10.010	24.935	-13.216
	247	LY4	CC	10.432	28.040	-14.098	247	LY4	CD1	10.094	29.331	-13.200
	247	LY4	CD2	11.924	27.921	-14.327	248	LY4	N	7.064	27.063	-14.632
	248	LY4	CA	6.404	28.033	-15.944	248	LY4	C	7.436	28.346	-17.065
40	248	LY4	CB	8.939	28.793	-16.912	248	LY4	CC	9.969	29.210	-19.099
	248	LY4	CC1	6.099	30.541	-15.952	248	LY4	CC2	4.263	28.929	-14.867
	248	LY4	CC2	8.399	31.765	-16.262	249	AS4	N	7.007	27.063	-18.233

269	ALA	CA	7.852	27.979	-21.439	269	ALA	C	7.852	27.979	-21.439
269	ALA	D	1.968	27.962	-21.402	269	ALA	CB	6.492	26.823	-21.401
269	ALA	CG	9.181	26.426	-21.216	269	ALA	CD1	6.993	27.636	-21.321
269	ALA	CD2	11.021	21.796	-21.073	272	ALA	D	1.968	26.962	-21.374
270	VAL	CA	6.819	27.118	-21.014	270	VAL	Z	6.819	26.967	-21.014
270	VAL	D	6.817	27.969	-21.072	270	VAL	CB	6.646	26.710	-21.072
270	VAL	CG1	6.869	22.717	-21.070	270	VAL	CG2	6.418	26.862	-21.072
271	GLN	C	7.325	26.761	-21.033	271	GLN	CA	7.623	26.750	-21.033
271	GLN	D	6.869	27.914	-21.033	271	GLN	D	6.213	27.906	-21.033
271	GLN	CG	9.104	26.220	-21.064	271	GLN	CG	9.086	26.010	-21.064
271	GLN	CD	10.901	26.815	-21.082	271	GLN	CD1	11.969	26.240	-21.082
271	GLN	CD2	11.702	26.813	-21.010	272	ALA	D	1.977	26.996	-21.002
272	ALA	CA	6.244	26.722	-21.060	272	ALA	C	7.791	26.980	-21.060
272	ALA	D	2.891	21.601	-21.101	272	ALA	CB	6.763	26.762	-21.102
272	ALA	D	6.247	26.041	-21.101	272	ALA	CA	2.900	26.983	-21.060
272	ALA	C	2.891	27.029	-21.020	272	ALA	D	1.979	26.989	-21.060
272	ALA	CB	2.716	27.773	-21.023	272	ALA	D	1.978	26.964	-21.060
272	ALA	CB	2.993	26.941	-21.010	274	ALA	CA	2.109	26.766	-21.047
274	ALA	C	2.730	21.917	-21.090	274	ALA	D	9.920	26.769	-21.021
275	GLN	D	2.810	27.194	-21.014	275	GLN	CA	2.948	26.990	-21.027
275	GLN	C	1.101	27.201	-21.077	275	GLN	D	1.740	27.007	-21.010
275	GLN	D	1.133	27.341	-21.000	275	GLN	CB	9.666	26.796	-21.020
275	GLN	CB	2.851	26.644	-21.047	275	GLN	CB	-3.073	26.936	-21.030
275	GLN	CD1	-1.374	23.009	-21.020	275	GLN	CD2	-1.973	23.411	-21.030

The above structural studies together with the kinetic data presented herein and elsewhere (Philipp, M., et al. (1983) *Mol. Cell. Biochem.* 51, 5-32; Svendsen, I.B. (1976) *Carlsberg Res. Comm.* 41, 237-291; Markland, S.F. Id; Stauffe, D.C., et al. (1965) *J. Biol. Chem.* 244, 5333-5338) indicate that the subsites in the binding cleft of subtilisin are capable of interacting with substrate amino acid residues from P-4 to P-2'.

The most extensively studied of the above residues are Gly166, Gly169 and Ala152. These amino acids were identified as residues within the S-1 subsite. As seen in Fig. 3, which is a stereoview of the S-1 subsite, Gly166 and Gly169 occupy positions at the bottom of the S-1 subsite, whereas Ala152 occupies a position near the top of S-1, close to the catalytic Ser221.

All 19 amino acid substitutions of Gly166 and Gly169 have been made. As will be indicated in the examples which follow, the preferred replacement amino acids for Gly166 and/or Gly169 will depend on the specific amino acid occupying the P-1 position of a given substrate.

The only substitutions of Ala152 presently made and analyzed comprise the replacement of Ala152 with Gly and Ser. The results of these substitutions on P-1 specificity will be presented in the examples.

In addition to those residues specifically associated with specificity for the P-1 substrate amino acid, Tyr104 has been identified as being involved with P-4 specificity. Substitutions at Phe189 and Tyr217, however, are expected to respectively effect P-2' and P-1' specificity.

The catalytic activity of subtilisin has also been modified by single amino acid substitutions at Asn155. The catalytic triad of subtilisin is shown in Fig. 4. As can be seen, Ser221, His64 and Asp32 are positioned to facilitate nucleophilic attack by the serine hydroxylate on the carbonyl of the scissile peptide bond. Crystallographic studies of subtilisin (Robertus, et al. (1972) *Biochem.* 11, 4293-4303; Matthews, et al. (1975) *J. Biol. Chem.* 250, 7120-7126; Poulos, et al. (1976) *J. Biol. Chem.* 250, 1097-1103) show that two hydrogen bonds are formed with the oxyanion of the substrate transition state. One hydrogen bond donor is from the catalytic serine-221 main-chain amide while the other is from one of the NE2 protons of the asparagine-155 side chain. See Fig. 4.

Asn155 was substituted with Ala, Asp, His, Glu and Thr. These substitutions were made to investigate the stabilization of the charged tetrahedral intermediate of the transition state complex by the potential hydrogen bond between the side chain of Asn155 and the oxyanion of the intermediate. These particular substitutions caused large decreases in substrate turnover, kcat (200 to 4,000 fold), marginal decreases in substrate binding Km (up to 7 fold), and a loss in transition state stabilization energy of 2.2 to 4.7 kcal/mol. The retention of Km and the drop in kcat will make these mutant enzymes useful as binding proteins for specific peptide sequences, the nature of which will be determined by the specificity of the precursor protease.

Various other amino acid residues have been identified which affect alkaline stability. In some cases, mutants having altered alkaline stability also have altered thermal stability.

In *B. amyloliquefaciens* subtilisin residues Asp36, Ile107, Lys170, Ser204 and Lys213 have been identified as residues which upon substitution with a different amino acid alter the alkaline stability of the mutated enzyme as compared to the precursor enzyme. The substitution of Asp36 with Ala and the substitution of Lys170 with Glu each resulted in a mutant enzyme having a lower alkaline stability as compared to the wild type subtilisin. When Ile107 was substituted with Val, Ser204 substituted with Cys, Arg or Leu or Lys213 substituted with Arg, the mutant subtilisin had a greater alkaline stability as compared

to the wild type subtilisin. However, the mutant Ser204P demonstrated a decrease in alkaline stability.

In addition, other residues, identified as being associated with the modification of other properties of subtilisin, also affect alkaline stability. These residues include Ser24, Met50, Glu156, Gly166, Gly169 and Tyr217. Specifically the following particular substitutions result in an increased alkaline stability: Ser24C, Met50F, Gly156Q or S, Gly166A, H, K, N or Q, Gly169S or A, and Tyr217F, K, R or L. The mutant Met50V, on the other hand, results in a decrease in the alkaline stability of the mutant subtilisin as compared to wild type subtilisin.

Other residues involved in alkaline stability based on the alkaline stability screen include Asp197 and Met222. Particular mutants include Asp197(R or A) and Met 222 (all other amino acids).

Various other residues have been identified as being involved in thermal stability as determined by the thermal stability screen herein. These residues include the above identified residues which effect alkaline stability and Met199 and Tyr21. These latter two residues are also believed to be important for alkaline stability. Mutants at these residues include I199 and F21.

The amino acid sequence of *B. amyloliquefaciens* subtilisin has also been modified by substituting two or more amino acids of the wild-type sequence. Six categories of multiply substituted mutant subtilisin have been identified. The first two categories comprise thermally and oxidatively stable mutants. The next three other categories comprise mutants which combine the useful properties of any of several single mutations of *B. amyloliquefaciens* subtilisin. The last category comprises mutants which have modified alkaline and/or thermal stability.

The first category comprises double mutants in which two cysteine residues have been substituted at various amino acid residue positions within the subtilisin molecule. Formation of disulfide bridges between the two substituted cysteine residues results in mutant subtilisins with altered thermal stability and catalytic activity. These mutants include A21/C22/C87 and C24/C87 which will be described in more detail in Example 11.

The second category of multiple subtilisin mutants comprises mutants which are stable in the presence of various oxidizing agents such as hydrogen peroxide or peracids. Examples 1 and 2 describe these mutants which include F50/I124/Q222, F50/I124, F50/Q222, F50/L124/Q222, I124/Q222 and L124/Q222.

The third category of multiple subtilisin mutants comprises mutants with substitutions at position 222 combined with various substitutions at positions 166 or 169. These mutants, for example, combine the property of oxidative stability of the A222 mutation with the altered substrate specificity of the various 166 or 169 substitutions. Such multiple mutants include A166/A222, A166/C222, F166/C222, K166/A222, K166/C222, V166/A222 and V166/C222. The K166/A222 mutant subtilisin, for example, has a kcat/Km ratio which is approximately two times greater than that of the single A222 mutant subtilisin when compared using a substrate with phenylalanine as the P-1 amino acid. This category of multiple mutant is described in more detail in Example 12.

The fourth category of multiple mutants combines substitutions at position 156 (Glu to Q or S) with the substitution of Lys at position 166. Either of these single mutations improve enzyme performance upon substrates with glutamate as the P-1 amino acid. When these single mutations are combined, the resulting multiple enzyme mutants perform better than either precursor. See Example 9.

The fifth category of multiple mutants contain the substitution of up to four amino acids of the *B. amyloliquefaciens* subtilisin sequence. These mutants have specific properties which are virtually identical to the properties of the subtilisin from *B. licheniformis*. The subtilisin from *B. licheniformis* differs from *B. amyloliquefaciens* subtilisin at 87 out of 275 amino acids. The multiple mutant F50/S156/A169/L217 was found to have similar substrate specificity and kinetics to the licheniformis enzyme. (See Example 13.) However, this is probably due to only three of the mutations (S156, A169 and L217) which are present in the substrate binding region of the enzyme. It is quite surprising that, by making only three changes out of the 87 different amino acids between the sequence of the two enzymes, the *B. amyloliquefaciens* enzyme was converted into an enzyme with properties similar to *B. licheniformis* enzyme. Other enzymes in this series include F50/Q156/N166/L217 and F50/S156/L217.

The sixth category of multiple mutants includes the combination of substitutions at position 107 (Ile to V) with the substitution of Lys at position 213 with Arg, and the combination of substitutions of position 204 (preferably Ser to C or L but also to all other amino acids) with the substitution of Lys at position 213 with R. Other multiple mutants which have altered alkaline stability include Q156/K166, Q156/N166, S156/K166, S156/N166 (previously identified as having altered substrate specificity), and F50/S156/A169/L217 (previously identified as a mutant of *B. amyloliquefaciens* subtilisin having properties similar to subtilisin from *B. licheniformis*). The mutant F50/V107/R213 was constructed based on the observed increase in alkaline stability for the single mutants F50, V107 and R213. It was determined that the V107/R213 mutant had an increased alkaline stability as compared to the wild type subtilisin. In this particular mutant, the increased

alkaline stability was the result of the cumulative stability of each of the individual mutations. Similarly, the mutant F50/V107/R213 had an even greater alkaline stability as compared to the V107/R213 mutant indicating that the increase in the alkaline stability due to the F50 mutation was also cumulative.

Table IV summarizes the multiple mutants which have been made including those not mentioned above.

5 In addition, based in part on the above results, substitution at the following residues in subtilisin is expected to produce a multiple mutant having increased thermal and alkaline stability: Ser24, Met50, Ile107, Glu156, Gly166, Gly169, Ser204, Lys213, Gly215, and Tyr217.

TABLE IV

10

15

20

25

30

35

40

Double Mutants	Triple, Quadruple or Other Multiple
C22/C87	F50/I124/Q222
C24/C87	F50/L124/Q222
V45/V48	F50/L124/A222
C49/C94	A21/C22/C87
C49/C95	F50/S156/N166/L217
C50/C95	F50/Q156/N166/L217
C50/C110	F50/S156/A169/L217
F50/I124	F50/S156/L217
F50/Q222	F50/Q156/K166/L217
I124/Q222	F50/S156/K166/L217
Q156/D166	F50/Q156/K166/K217
Q156/K166	F50/S156/K166/K217
Q156/N166	F50/V107/R213
S156/D166	[S153/S156/A158/G159/S160/ Δ 161-164/I165/S166/A169/R170]
S156/K166	
S156/N166	L204/R213
S156/A169	R213/204A, E, Q, D, N, G, K, V, R, T, P, I, M, F, Y, W or H
A166/A222	
A166/C222	
F166/A222	V107/R213
F166/C222	
K166/A222	
K166/C222	
V166/A222	
V166/C222	
A169/A222	
A169/A222	
A169/C222	
A21/C22	

45

In addition to the above identified amino acid residues, other amino acid residues of subtilisin are also considered to be important with regard to substrate specificity. Mutation of each of these residues is expected to produce changes in the substrate specificity of subtilisin. Moreover, multiple mutations among these residues and among the previously identified residues are also expected to produce subtilisin mutants having novel substrate specificity.

50

Particularly important residues are His67, Ile107, Leu126 and Leu135. Mutation of His67 should alter the S-1' subsite, thereby altering the specificity of the mutant for the P-1' substrate residue. Changes at this position could also affect the pH activity profile of the mutant. This residue was identified based on the inventor's substrate modeling from product inhibitor complexes.

55

Ile107 is involved in P-4 binding. Mutation at this position thus should alter specificity for the P-4 substrate residue in addition to the observed effect on alkaline stability. Ile107 was also identified by molecular modeling from product inhibitor complexes.

The S-2 binding site includes the Leu126 residue. Modification at this position should therefore affect P-2 specificity. Moreover, this residue is believed to be important to convert subtilisin to an amino peptidase.

The pH activity profile should also be modified by appropriate substitution. These residues were identified from inspection of the refined model, the three dimensional structure from modeling studies. A longer side chain is expected to preclude binding of any side chain at the S-2 subsite. Therefore, binding would be restricted to subsites S-1, S-1', S-2', S-3' and cleavage would be forced to occur after the amino terminal peptide.

Leu135 is in the S-4 subsite and if mutated should alter substrate specificity for P-4 if mutated. This residue was identified by inspection of the three-dimensional structure and modeling based on the product inhibitor complex of F222.

In addition to these sites, specific amino acid residues within the segments 97-103, 126-129 and 213-215 are also believed to be important to substrate binding.

Segments 97-103 and 126-129 form an antiparallel beta sheet with the main chain of substrate residues P-4 through P-2. Mutating residues in those regions should affect the substrate orientation through main chain (enzyme) - main chain (substrate) interactions, since the main chain of these substrate residues do not interact with these particular residues within the S-4 through S-2 subsites.

Within the segment 97-103, Gly97 and Asp99 may be mutated to alter the position of residues 101-103 within the segment. Changes at these sites must be compatible, however. In *B. amyloliquefaciens* subtilisin Asp99 stabilizes a turn in the main chain tertiary folding that affects the direction of residues 101-103. *B. licheniformis* subtilisin Asp97, functions in an analogous manner.

In addition to Gly97 and Asp99, Ser101 interacts with Asp99 in *B. amyloliquefaciens* subtilisin to stabilize the same main chain turn. Alterations at this residue should alter the 101-103 main chain direction. Mutations at Glu103 are also expected to affect the 101-103 main chain direction.

The side chain of Gly102 interacts with the substrate P-3 amino acid. Side chains of substituted amino acids thus are expected to significantly affect specificity for the P-3 substrate amino acids.

All the amino acids within the 127-129 segment are considered important to substrate specificity. Gly127 is positioned such that its side chain interacts with the S-1 and S-3 subsites. Altering this residue thus should alter the specificity for P-1 and P-3 residues of the substrate.

The side chain of Gly128 comprises a part of both the S-2 and S-4 subsites. Altered specificity for P-2 and P-4 therefore would be expected upon mutation. Moreover, such mutation may convert subtilisin into an amino peptidase for the same reasons substitutions of Leu126 would be expected to produce that result.

The Pro129 residue is likely to restrict the conformational freedom of the sequence 126-133, residues which may play a major role in determining P-1 specificity. Replacing Pro may introduce more flexibility thereby broadening the range of binding capabilities of such mutants.

The side chain of Lys213 is located within the S-3 subsite. All of the amino acids within the 213-215 segment are also considered to be important to substrate specificity. Accordingly, altered P-3 substrate specificity is expected upon mutation of this residue.

The Tyr214 residue does not interact with substrate but is positioned such that it could affect the conformation of the hair pin loop 204-217.

Finally, mutation of the Gly215 residue should affect the S-3' subsite, and thereby alter P-3' specificity.

In addition to the above substitutions of amino acids, the insertion or deletion of one or more amino acids within the external loop comprising residues 152-172 may also affect specificity. This is because these residues may play a role in the "secondary contact region" described in the model of streptomyces subtilisin inhibitor complexed with subtilisin. Hirano, et al. (1984) *J. Mol. Biol.* 178, 389-413. Thermitase K has a deletion in this region, which eliminates several of these "secondary contact" residues. In particular, deletion of residues 161 through 164 is expected to produce a mutant subtilisin having modified substrate specificity. In addition, a rearrangement in this area induced by the deletion should alter the position of many residues involved in substrate binding, predominantly at P-1. This, in turn, should affect overall activity against proteinaceous substrates

The effect of deletion of residues 161 through 164 has been shown by comparing the activity of the wild type (WT) enzyme with a mutant enzyme containing this deletion as well as multiple substitutions (i.e., S153/S156/A158/G159/S160/Δ161-164/I165/S166/A169/R170). This produced the following results:

TABLE V

	kcat	Km	kcat/Km
WT	50	1.4×10^{-4}	3.6×10^5
Deletion mutant	8	5.0×10^{-6}	1.6×10^6

The WT has a k_{cat} 6 times greater than the deletion mutant but substrate binding is 28 fold tighter by the deletion mutant. The overall efficiency of the deletion mutant is thus 4.4 times higher than the WT enzyme.

All of these above identified residues which have yet to be substituted, deleted or inserted into are presented in Table VI.

TABLE VI

Substitution/Insertion/Deletion	
Residues	
His67	Ala152
Leu126	Ala153
Leu135	Gly154
Gly97	Asn155
Asp99	Gly156
Ser101	Gly157
Gly102	Gly160
Glu103	Thr158
Leu126	Ser159
Gly127	Ser161
Gly128	Ser162
Pro129	Ser163
Tyr214	Thr164
Gly215	Val165
Gly166	Gly169
Tyr167	Lys170
Pro168	Tyr171
	Pro172

The following disclosure is intended to serve as a representation of embodiments herein, and should not be construed as limiting the scope of this application. These specific examples disclose the construction of certain of the above identified mutants. The construction of the other mutants, however, is apparent from the disclosure herein and that presented in EPO Publication No. 0130756.

All literature citations are expressly incorporated by reference.

EXAMPLE 1

Identification of Peracid Oxidizable Residues of Subtilisin Q222 and L222

As shown in Figures 6A and 6B, organic peracid oxidants inactivate the mutant subtilisins Met222L and Met222Q (L222 and Q222). This example describes the identification of peracid oxidizable sites in these mutant subtilisins.

First, the type of amino acid involved in peracid oxidation was determined. Except under drastic conditions (Means, G.E., et al. (1971) Chemical Modifications of Proteins, Holden-Day, S.F., CA, pp. 160-162), organic peracids modify only methionine and tryptophan in subtilisin. Difference spectra of the enzyme over the 250nm to 350nm range were determined during an inactivation titration employing the reagent, diperdodecanoic acid (DPDA) as oxidant. Despite quantitative inactivation of the enzyme, no change in absorbance over this wavelength range was noted as shown in Figures 7A and 7B indicating that tryptophan was not oxidized. Fontana, A., et al. (1980) Methods in Peptide and Protein Sequence Analysis - (C. Birr ed.) Elsevier, New York, p. 309. The absence of tryptophan modification implied oxidation of one or more of the remaining methionines of *B. amyloliquefaciens* subtilisin. See Figure 1.

To confirm this result the recombinant subtilisin Met222F was cleaved with cyanogen bromide (CNBr) both before and after oxidation by DPDA. The peptides produced by CNBr cleavage were analyzed on high resolution SDS-pyridine peptide gels (SPG).

Subtilisin Met222F (F222) was oxidized in the following manner. Purified F222 was resuspended in 0.1 M sodium borate pH 9.5 at 10 mg/ml and was added to a final concentration of 26 diperdodecanoic acid

(DPDA) at 26 mg/ml was added to produce an effective active oxygen concentration of 30 ppm. The sample was incubated for at least 30 minutes at room temperature and then quenched with 0.1 volume of 1 M Tris pH 8.6 buffer to produce a final concentration of 0.1 M Tris pH 8.6). 3mM phenylmethylsulfonyl fluoride (PMSF) was added and 2.5 ml of the sample was applied to a Pharmacia PD10 column equilibrated in 10 mM sodium phosphate pH 6.2, 1 mM PMSF. 3.5 ml of 10 mM sodium phosphate pH6.2, 1mM PMSF was applied and the eluant collected.

F222 and DPDA oxidized F222 were precipitated with 9 volumes of acetone at -20 °C. The samples were resuspended at 10 mg/ml in 8M urea in 88% formic acid and allowed to sit for 5 minutes. An equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein) and the samples incubated for 2 hours at room temperature in the dark. Prior to gel electrophoresis, the samples were lyophilized and resuspended at 2-5 mg/ml in sample buffer (1% pyridine, 5% NaDodSO₄, 5% glycerol and bromophenol blue) and disassociated at 95 °C for 3 minutes.

The samples were electrophoresed on discontinuous polyacrylamide gels (Kyte, J., et al. (1953) *Anal. Bioch.* 133, 515-522). The gels were stained using the Pharmacia silver staining technique (Sammons, D.W., et al. (1981) *Electrophoresis* 2 135-141).

The results of this experiment are shown in Figure 8. As can be seen, F222 treated with CNBr only gives nine resolved bands on SPG. However, when F222 is also treated with DPDA prior to cleavage, bands X, 7 and 9 disappear whereas bands 5 and 6 are greatly increased in intensity.

In order to determine which of the methionines were effected, each of the CNBr peptides was isolated by reversed phase HPLC and further characterized. The buffer system in both Solvent A (aqueous) and Solvent B (organic) for all HPLC separations was 0.05% triethylamine/trifluoroacetic acid (TEA-TFA). In all cases unless noted, solvent A consisted of 0.05% TEA-TFA in H₂O, solvent B was 0.05% TEA-TFA in 1-propanol, and the flow rate was 0.5 ml/minute.

For HPLC analysis, two injections of 1 mg enzyme digest were used. Three samples were acetone precipitated, washed and dried. The dried 1 mg samples were resuspended at 10 mg/ml in 8M urea, 88% formic acid; an equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein). After incubation for 2 hours in the dark at room temperature, the samples were desalted on a 0.8 cm X 7 cm column of Tris Acryl GF05 coarse resin (IBF, Paris, France) equilibrated with 40% solvent B, 60% solvent A. 200 ul samples were applied at a flow rate of 1 ml a minute and 1.0-1.2 ml collected by monitoring the absorbance at 280nm. Prior to injection on the HPLC, each desalted sample was diluted with 3 volumes of solvent A. The samples were injected at 1.0 ml/min (2 minutes) and the flow then adjusted to 0.5 ml/min (100% A). After 2 minutes, a linear gradient to 60% B at 1.0% B/min was initiated. From each 1 mg run, the pooled peaks were sampled (50ul) and analyzed by gel electrophoresis as described above.

Each polypeptide isolated by reversed phase HPLC was further analyzed for homogeneity by SPG. The position of each peptide on the known gene sequence (Wells, J.A., et al. (1983) *Nucleic Acids Res.* 11 7911-7924) was obtained through a combination of amino acid compositional analysis and, where needed, amino terminal sequencing.

Prior to such analysis the following peptides were to rechromatographed.

1. CNBr peptides from F222 not treated with DPDA:

Peptide 5 was subjected to two additional reversed phase separations. The 10 cm C4 column was equilibrated to 80%A/ 20%B and the pooled sample applied and washed for 2 minutes. Next an 0.5% ml B/min gradient was initiated. Fractions from this separation were again rerun, this time on the 25 cm C4 column, and employing 0.05% TEA-TFA in acetonitrile/1-propanol (1:1) for solvent B. The gradient was identical to the one just described.

Peptide "X" was subjected to one additional separation after the initial chromatography. The sample was applied and washed for 2 minutes at 0.5ml/min (100%A), and a 0.5% ml B/min gradient was initiated.

Peptides 7 and 9 were rechromatographed in a similar manner to the first rerun of peptide 5.

Peptide 8 was purified to homogeneity after the initial separation.

2. CNBr Peptides from DPDA Oxidized F222:

Peptides 5 and 6 from a CNBr digest of the oxidized F222 were purified in the same manner as peptide 5 from the untreated enzyme.

Amino acid compositional analysis was obtained as follows. Samples (~1nM each amino acid) were dried, hydrolyzed in vacuo with 100 ul 6N HCl at 106 °C for 24 hours and then dried in a Speed Vac. The samples were analyzed on a Beckmann 6300 AA analyzer employing ninhydrin detection.

Amino terminal sequence data was obtained as previously described (Rodriguez, H., et al. (1984) Anal. Biochem. 134, 538-547).

The results are shown in Table VII and Figure 9.

TABLE VII

Amino and COOH termini of CNBr fragments Terminus and Method		
Fragment	amino, method	COOH, method
X	1, sequence	50, composition
9	51, sequence	119, composition
7	125, sequence	199, composition
8	200, sequence	275, composition
5ox	1, sequence	119, composition
6ox	120, composition	199, composition

Peptides 5ox and 6ox refer to peptides 5 and 6 isolated from CNBr digests of the oxidized protein where their respective levels are enhanced.

From the data in Table VII and the comparison of SPG tracks for the oxidized and native protein digests in Figure 8, it is apparent that (1) Met50 is oxidized leading to the loss of peptides X and 9 and the appearance of 5; and (2) Met124 is also oxidized leading to the loss of peptide 7 and the accumulation of peptide 6. Thus oxidation of *B. amyloquifaciens* subtilisin with the peracid, diperdocecanoic acid leads to the specific oxidation of methionine at residues 50 and 124.

EXAMPLE 2

Substitution at Met50 and Met124 in Subtilisin Met222Q

The choice of amino acid for substitution at Met50 was based on the available sequence data for subtilisins from *B. licheniformis* (Smith, E.C., et al. (1968) *J. Biol. Chem.* 243, 2184-2191), *B. DY* (Nedkov, P., et al. (1983) *Hoppe Saylor's Z. Physiol. Chem.* 364 1537-1540), *B. amylosacchariticus* (Markland, F.S., et al. (1967) *J. Biol. Chem.* 242 5198-5211) and *B. subtilis* (Stahl, M.L., et al. (1984) *J. Bacteriol.* 158, 411-418). In all cases, position 50 is a phenylalanine. See Figure 5. Therefore, Phe50 was chosen for construction.

At position 124, all known subtilisins possess a methionine. See Figure 5. Molecular modelling of the x-ray derived protein structure was therefore rehired to determine the most probable candidates for substitution. From all 19 candidates, isoleucine and leucine were chosen as the best residues to employ. In order to test whether or not modification at one site but not both was sufficient to increase oxidative stability, all possible combinations were built on the Q222 backbone (F50/Q222, I124/Q222, F50/I124/Q222).

A. Construction of Mutations Between Codons 45 and 50

All manipulations for cassette mutagenesis were carried out on pS4.5 using methods disclosed in EPO Publication No. 0130756 and Wells, J.A., et al. (1985) *Gene* 34, 315-323. The pΔ50 in Fig. 10, line 4, mutations was produced using the mutagenesis primer shown in Fig. 10, line 6, and employed an approach designated as restriction-purification which is described below. Briefly, a M13 template containing the subtilisin gene, M13mp11-SUBT was used for heteroduplex synthesis (Adelman, et al (1983), *DNA* 2, 183-193). Following transfection of JM101 (ATCC 33876), the 1.5 kb *EcoRI*-*Bam*HI fragment containing the subtilisin gene was subcloned from M13mp11 SUBT rf into a recipient vector fragment of pBS42 the construction of which is described in EPO Publication No. 0130756. To enrich for the mutant sequence (pΔ50, line 4), the resulting plasmid pool was digested with *Kpn*I, and linear molecules were purified by polyacrylamide gel electrophoresis. Linear molecules were ligated back to a circular form, and transformed into *E. coli* MM294 cells (ATCC 31446). Isolated plasmids were screened by restriction analysis for the *Kpn*I site. *Kpn*I⁺ plasmids were sequenced and confirmed the pΔ50 sequence. Asterisks in Figure 11 indicate the bases that are mutated from the wild type sequence (line 4). pΔ50 (line 4) was cut with *Stu*I and *EcoRI* and the 0.5 Kb fragment containing the 5' half of the subtilisin gene was purified (fragment 1). pΔ50 (line 4) was digested with *Kpn*I and *EcoRI* and the 4.0 Kb fragment containing the 3' half of the subtilisin gene and vector sequences was purified (fragment 2). Fragments 1 and 2 (line 5), and duplex DNA

cassettes coding for mutations desired (shaded sequence, line 6) were mixed in a molar ratio of 1:1:10, respectively. For the particular construction of this example the DNA cassette contained the triplet TTT for codon 50 which encodes Phe. This plasmid was designated pF50. The mutant subtilisin was designated F50.

B. Construction of Mutation Between Codons 122 and 127

The procedure of Example 2A was followed in substantial detail except that the mutagenesis primer of Figure 11, line 7 was used and restriction-purification for the *EcoRV* site in p Δ 124 was used. In addition, the DNA cassette (shaded sequence, Figure 11, line 6) contained the triplet ATT for codon 124 which encodes Ile and CTT for Leu. Those plasmids which contained the substitution of Ile for Met124 were designated pI124. The mutant subtilisin was designated I124.

C. Construction of Various F50/I124/Q222 Multiple Mutants

The triple mutant, F50/I124/Q222, was constructed from a three-way ligation in which each fragment contained one of the three mutations. The single mutant Q222 (pQ222) was prepared by cassette mutagenesis as described in EPO Publication No. 0130756. The F50 mutation was contained on a 2.2kb *Avall* to *PvuII* fragment from pF50; the I124 mutation was contained on a 260 bp *PvuII* to *Avall* fragment from pI124; and the Q222 mutation was contained on 2.7 kb *Avall* to *Avall* fragment from pQ222. The three fragments were ligated together and transformed into *E. coli* MM294 cells. Restriction analysis of plasmids from isolated transformants confirmed the construction. To analyze the final construction it was convenient that the *Avall* site at position 798 in the wild-type subtilisin gene was eliminated by the I124 construction.

The F50/Q222 and I124/Q222 mutants were constructed in a similar manner except that the appropriate fragment from pS4.5 was used for the final construction.

D. Oxidative Stability of Q222 Mutants

The above mutants were analyzed for stability to peracid oxidation. As shown in Fig. 12, upon incubation with dperdodecanoic acid (protein 2mg/mL, oxidant 75ppm[O]), both the I124/Q222 and the F50/I124/Q222 are completely stable whereas the F50/Q222 and the Q222 are inactivated. This indicates that conversion of Met124 to I124 in subtilisin Q222 is sufficient to confer resistance to organic peracid oxidants.

EXAMPLE 3

Subtilisin Mutants Having Altered Substrate Specificity-Hydrophobic Substitutions at Residues 166

Subtilisin contains an extended binding cleft which is hydrophobic in character. A conserved glycine at residue 166 was replaced with twelve non-ionic amino acids which can project their side-chains into the S-1 subsite. These mutants were constructed to determine the effect of changes in size and hydrophobicity on the binding of various substrates.

A. Kinetics for Hydrolysis of Substrates Having Altered P-1 Amino Acids by Subtilisin from *B. Amylolyquefaciens*

Wild-type subtilisin was purified from *B. subtilis* culture supernatants expressing the *B. amylolyquefaciens* subtilisin gene (Wells, J.A., et al. (1983) *Nucleic Acids Res.* 11, 7911-7925) as previously described (Estell, D.A., et al. (1985) *J. Biol. Chem.* 260, 6518-6521). Details of the synthesis of tetrapeptide substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide (where X is the P1 amino acid) are described by DelMar, E.G., et al. (1979) *Anal. Biochem.* 99, 316-320. Kinetic parameters, K_m (M) and k_{cat} (s⁻¹) were measured using a modified progress curve analysis (Estell, D.A., et al. (1985) *J. Biol. Chem.* 260, 6518-6521). Briefly, plots of rate versus product concentration were fit to the differential form of the rate equation using a non-linear regression algorithm. Errors in k_{cat} and K_m for all values reported are less than five percent. The various substrates in Table VIII are ranged in order of decreasing hydrophobicity. Nozaki, Y. (1971), *J. Biol. Chem.* 246, 2211-2217; Tanford C. (1978) *Science* 200, 1012).

TABLE VIII

P1 substrate Amino Acid	kcat(S ⁻¹)	1/Km(M ⁻¹)	kcat/Km (s ⁻¹ M ⁻¹)
Phe	50	7,100	360,000
Tyr	28	40,000	1,100,000
Leu	24	3,100	75,000
Met	13	9,400	120,000
His	7.9	1,600	13,000
Ala	1.9	5,500	11,000
Gly	0.003	8,300	21
Gln	3.2	2,200	7,100
Ser	2.8	1,500	4,200
Glu	0.54	32	16

The ratio of kcat/Km (also referred to as catalytic efficiency) is the apparent second order rate constant for the conversion of free enzyme plus substrate (E + S) to enzyme plus products (E + P) (Jencks, W.P., *Catalysis in Chemistry and Enzymology* (McGraw-Hill, 1969) pp. 321-436; Fersht, A., *Enzyme Structure and Mechanism* (Freeman, San Francisco, 1977) pp. 226-287). The log (kcat/Km) is proportional to transition state binding energy, ΔG^\ddagger . A plot of the log kcat/Km versus the hydrophobicity of the P1 side-chain (Figure 14) shows a strong correlation ($r = 0.98$), with the exception of the glycine substrate which shows evidence for non-productive binding. These data show that relative differences between transition-state binding energies can be accounted for by differences in P-1 side-chain hydrophobicity. When the transition-state binding energies are calculated for these substrates and plotted versus their respective side-chain hydrophobicities, the line slope is 1.2 (not shown). A slope greater than unity, as is also the case for chymotrypsin (Fersht, A., *Enzyme Structure and Mechanism* (Freeman, San Francisco, 1977) pp. 226-287; Harper, J.W., et al. (1984) *Biochemistry*, 23, 2995-3002), suggests that the P1 binding cleft is more hydrophobic than ethanol or dioxane solvents that were used to empirically determine the hydrophobicity of amino acids (Nozaki, Y., et al. *J. Biol. Chem.* (1971) 246, 2211-2217; Tanford, C. (1978) *Science* 200, 1012).

For amide hydrolysis by subtilisin, kcat can be interpreted as the acylation rate constant and Km as the dissociation constant, for the Michaelis complex (E•S), Ks. Gutfreund, H., et al (1956) *Biochem. J.* 63, 656. The fact that the log kcat, as well as log 1/Km, correlates with substrate hydrophobicity is consistent with proposals (Robertus, J.D., et al. (1972) *Biochemistry* 11, 2439-2449; Robertus, J.D., et al. (1972) *Biochemistry* 11, 4293-4303) that during the acylation step the P-1 side-chain moves deeper into the hydrophobic cleft as the substrate advances from the Michaelis complex (E•S) to the tetrahedral transition-state complex (E•S^{*}). However, these data can also be interpreted as the hydrophobicity of the P1 side-chain effecting the orientation, and thus the susceptibility of the scissile peptide bond to nucleophilic attack by the hydroxyl group of the catalytic Ser221.

The dependence of kcat/Km on P-1 side chain hydrophobicity suggested that the kcat/Km for hydrophobic substrates may be increased by increasing the hydrophobicity of the S-1 binding subsite. To test this hypothesis, hydrophobic amino acid substitutions of Gly166 were produced.

Since hydrophobicity of aliphatic side-chains is directly proportional to side-chain surface area (Rose, G.D., et al. (1985) *Science* 229, 834-838; Reynolds, J.A., et al. (1974) *Proc. Natl. Acad. Sci. USA* 71, 2825-2927), increasing the hydrophobicity in the S-1 subsite may also sterically hinder binding of larger substrates. Because of difficulties in predicting the relative importance of these two opposing effects, we elected to generate twelve non-charged mutations at position 166 to determine the resulting specificities against non-charged substrates of varied size and hydrophobicity.

B. Cassette Mutagenesis of the P1 Binding Cleft

The preparation of mutant subtilisins containing the substitution of the hydrophobic amino acids Ala, Val and Phe into residue 166 has been described in EPO Publication No. 0130756. The same method was used to produce the remaining hydrophobic mutants at residue 166. In applying this method, two unique and silent restriction sites were introduced in the subtilisin genes to closely flank the target codon 166. As can be seen in Figure 13, the wild type sequence (line 1) was altered by site-directed mutagenesis in M13 using the indicated 37mer mutagenesis primer, to introduce a 13 bp deletion (dashedline) and unique SacI and XmaI sites (underlined sequences) that closely flank codon 166. The subtilisin gene fragment was subcloned back into the *E. coli* - *B. subtilis* shuttle plasmid, pBS42, giving the plasmid pΔ166 (Figure 13,

line 2). pΔ166 was cut open with SacI and XmaI, and gapped linear molecules were purified (Figure 13, line 3). Pools of synthetic oligonucleotides containing the mutation of interest were annealed to give duplex DNA cassettes that were ligated into gapped pΔ166 (underlined and overlined sequences in Figure 13, line 4). This construction restored the coding sequence except over position 166(NNN; line 4). Mutant sequences were confirmed by dideoxy sequencing. Asterisks denote sequence changes from the wild type sequence. Plasmids containing each mutant B. amyloliquefaciens subtilisin gene were expressed at roughly equivalent levels in a protease deficient strain of B. subtilis, BG2036 as previously described. EPO Publication No. 0130756; Yang, M., et al. (1984) J. Bacteriol. 160, 15-21; Estell, D.A., et al (1985) J. Biol. Chem. 260, 6518-6521.

C. Narrowing Substrate Specificity by Steric Hindrance

To probe the change in substrate specificity caused by steric alterations in the S-1 subsite, position 166 mutants were kinetically analyzed versus P1 substrates of increasing size (i.e., Ala, Met, Phe and Tyr). Ratios of k_{cat}/K_m are presented in log form in Figure 15 to allow direct comparisons of transition-state binding energies between various enzyme-substrate pairs.

According to transition state theory, the free energy difference between the free enzyme plus substrate ($E + S$) and the transition state complex ($E \cdot S^*$) can be calculated from equation (1),

$$(1) \quad \Delta G_T^\ddagger = -RT \ln k_{cat}/K_m + RT \ln kT/h$$

in which k_{cat} is the turnover number, K_m is the Michaelis constant, R is the gas constant, T is the temperature, k is Boltzmann's constant, and h is Planck's constant. Specificity differences are expressed quantitatively as differences between transition state binding energies (i.e., $\Delta\Delta G_T^\ddagger$), and can be calculated from equation (2).

$$(2) \quad \Delta\Delta G_T^\ddagger = -RT \ln (k_{cat}/K_m)_A / (k_{cat}/K_m)_B$$

A and B represent either two different substrates assayed against the same enzyme, or two mutant enzymes assayed against the same substrate.

As can be seen from Figure 15A, as the size of the side-chain at position 166 increases the substrate preference shifts from large to small P-1 side-chains. Enlarging the side-chain at position 166 causes k_{cat}/K_m to decrease in proportion to the size of the P-1 substrate side-chain (e.g., from Gly166 (wild-type) through W166, the k_{cat}/K_m for the Tyr substrate is decreased most followed in order by the Phe, Met and Ala P-1 substrates).

Specific steric changes in the position 166 side-chain, such as the presence of a β -hydroxyl group, β - or γ -aliphatic branching, cause large decreases in k_{cat}/K_m for larger P1 substrates. Introducing a β -hydroxyl group in going from A166 (Figure 15A) to S166 (Figure 15B), causes an 8 fold and 4 fold reduction in k_{cat}/K_m for Phe and Tyr substrates, respectively, while the values for Ala and Met substrates are unchanged. Producing a β -branched structure, in going from S166 to T166, results in a drop of 14 and 4 fold in k_{cat}/K_m for Phe and Tyr, respectively. These differences are slightly magnified for V166 which is slightly larger and isosteric with T166. Enlarging the β -branched substituents from V166 to I166 causes a lowering of k_{cat}/K_m between two and six fold toward Met, Phe and Tyr substrates. Inserting a γ -branched structure, by replacing M166 (Figure 15A) with L166 (Figure 15B), produces a 5 fold and 18 fold decrease in k_{cat}/K_m for Phe and Tyr substrates, respectively. Aliphatic γ -branching appears to induce less steric hindrance toward the Phe P-1 substrate than β -branching, as evidenced by the 100 fold decrease in k_{cat}/K_m for the Phe substrate in going from L166 to I166.

Reductions in k_{cat}/K_m resulting from increases in side chain size in the S-1 subsite, or specific structural features such as β - and γ -branching, are quantitatively illustrated in Figure 16. The k_{cat}/K_m values for the position 166 mutants determined for the Ala, Met, Phe, and Tyr P-1 substrates (top panel through bottom panel, respectively), are plotted versus the position 166 side-chain volumes (Chothia, C. (1984) Ann. Rev. Biochem. 53, 537-572). Catalytic efficiency for the Ala substrate reaches a maximum for

I166, and for the Met substrate it reaches a maximum between V166 and L166. The Phe substrate shows a broad k_{cat}/K_m peak but is optimal with A166. Here, the β -branched position 166 substitutions form a line that is parallel to, but roughly 50 fold lower in k_{cat}/K_m than side-chains of similar size [i.e., C166 versus T166, L166 versus I166]. The Tyr substrate is most efficiently utilized by wild type enzyme (Gly166), and there is a steady decrease as one proceeds to large position 166 side-chains. The β -branched and γ -branched substitutions form a parallel line below the other non-charged substitutions of similar molecular volume.

The optimal substitution at position 166 decreases in volume with increasing volume of the P1 substrate [i.e., I166/Ala substrate, L166/Met substrate, A166/Phe substrate, Gly166/Tyr substrate]. The combined volumes for these optimal pairs may approximate the volume for productive binding in the S-1 subsite. For the optimal pairs, Gly166/Tyr substrate, A166/Phe substrate, L166/Met substrate, V166/Met substrate, and I166/Ala substrate, the combined volumes are 266,295,313,339 and 261 Å³, respectively. Subtracting the volume of the peptide backbone from each pair (i.e., two times the volume of glycine), an average side-chain volume of 160±32 Å³ for productive binding can be calculated.

The effect of volume, in excess to the productive binding volume, on the drop in transition-state binding energy can be estimated from the Tyr substrate curve (bottom panel, Figure 16), because these data, and modeling studies (Figure 2), suggest that any substitution beyond glycine causes steric repulsion. A best-fit line drawn to all the data ($r = 0.87$) gives a slope indicating a loss of roughly 3 kcal/mol in transition state binding energy per 100 Å³ of excess volume. (100 Å³ is approximately the size of a leucyl side-chain.)

D. Enhanced Catalytic Efficiency Correlates with Increasing Hydrophobicity of the Position 166 Substitution

Substantial increases in k_{cat}/K_m occur with enlargement of the position 166 side-chain, except for the Tyr P-1 substrate (Figure 16). For example, k_{cat}/K_m increases in progressing from Gly166 to I166 for the Ala substrate (net of ten-fold), from Gly166 to L166 for the Met substrate (net of ten-fold) and from Gly166 to A166 for the Phe substrate (net of two-fold). The increases in k_{cat}/K_m cannot be entirely explained by the attractive terms in the van der Waals potential energy function because of their strong distance dependence ($1/r^6$) and because of the weak nature of these attractive forces (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Levitt, M. (1976) J. Mol. Biol. 104, 59-107). For example, Levitt (Levitt, M. (1976) J. Mol. Biol. 104, 59-107) has calculated that the van der Waals attraction between two methionyl residues would produce a maximal interaction energy of roughly -0.2 kcal/mol. This energy would translate to only 1.4 fold increase in k_{cat}/K_m .

The increases of catalytic efficiency caused by side-chain substitutions at position 166 are better accounted for by increases in the hydrophobicity of the S-1 subsite. The increase k_{cat}/K_m observed for the Ala and Met substrates with increasing position 166 side-chain size would be expected, because hydrophobicity is roughly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927).

Another example that can be interpreted as a hydrophobic effect is seen when comparing k_{cat}/K_m for isosteric substitutions that differ in hydrophobicity such as S166 and C166 (Figure 16). Cysteine is considerably more hydrophobic than serine (-1.0 versus +0.3 kcal/mol) (Nozaki, Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012). The difference in hydrophobicity correlates with the observation that C166 becomes more efficient relative to Ser166 as the hydrophobicity of the substrates increases (i.e., Ala < Met < Tyr < Phe). Steric hindrance cannot explain these differences because serine is considerably smaller than cysteine (99 versus 118 Å³). Paul, I.C., Chemistry of the -SH Group (ed. S. Patai, Wiley Interscience, New York, 1974) pp. 111-149.

E. Production of an Elastase-Like Specificity in Subtilisin

The I166 mutation illustrates particularly well that large changes in specificity can be produced by altering the structure and hydrophobicity of the S-1 subsite by a single mutation (Figure 17). Progressing through the small hydrophobic substrates, a maximal specificity improvement over wild type occurs for the Val substrate (16 fold in k_{cat}/K_m). As the substrate side chain size increases, these enhancements shrink to near unity (i.e., Leu and His substrates). The I166 enzyme becomes poorer against larger aromatic substrates of increasing size (e.g., I166 is over 1,000 fold worse against the Tyr substrate than is Gly166). We interpret the increase in catalytic efficiency toward the small hydrophobic substrates for I166 compared to Gly166 to the greater hydrophobicity of isoleucine (i.e., -1.8 kcal/mol versus 0). Nozaki, Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012. The decrease in catalytic efficiency

toward the very large substrates for I166 versus Gly166 is attributed to steric repulsion.

The specificity differences between Gly166 and I166 are similar to the specificity differences between chymotrypsin and the evolutionary relative, elastase (Harper, J.W., *et al* (1984) *Biochemistry* **23**, 2995-3002). In elastase, the bulky amino acids, Thr and Val, block access to the P-1 binding site for large hydrophobic substrates that are preferred by chymotrypsin. In addition, the catalytic efficiencies toward small hydrophobic substrates are greater for elastase than for chymotrypsin as we observe for I166 versus Gly166 in subtilisin.

EXAMPLE 4

Substitution of Ionic Amino Acids for Gly166

The construction of subtilisin mutants containing the substitution of the ionic amino acids Asp, Asn, Gln, Lys and Arg are disclosed in EPO Publication No. 0130756. The present example describes the construction of the mutant subtilisin containing Glu at position 166 (E166) and presents substrate specificity data on these mutants. Further data on position 166 and 156 single and double mutants is presented *infra*.

p Δ 166, described in Example 3, was digested with SacI and XmaI. The double strand DNA cassette (underlined and overlined) of line 4 in Figure 13 contained the triplet GAA for the codon 166 to encode the replacement of Glu for Gly166. This mutant plasmid designated pQ166 was propagated in BG2036 as described. This mutant subtilisin, together with the other mutants containing ionic substituent amino acids at residue 166, were isolated as described and further analyzed for variations in substrate specificity.

Each of these mutants was analyzed with the tetrapeptide substrates, succinyl-L-AlaL-AlaProL-X-p-nitroanilide, where X was Phe, Ala and Glu.

The results of this analysis are shown in Table IX.

TABLE IX

Position 166	P-1 Substrate (kcat/Km x 10 ⁻⁴)		
	Phe	Ala	Glu
Gly (wild type)	36.0	1.4	0.002
Asp (D)	0.5	0.4	<0.001
Glu (E)	3.5	0.4	<0.001
Asn (N)	18.0	1.2	0.004
Gln (Q)	57.0	2.6	0.002
Lys (K)	52.0	2.8	1.2
Arg (R)	42.0	5.0	0.08

These results indicate that charged amino acid substitutions at Gly166 have improved catalytic efficiencies (kcat/Km) for oppositely charged P-1 substrates (as much as 500 fold) and poorer catalytic efficiency for like charged P-1 substrates.

EXAMPLE 5

Substitution of Glycine at Position 169

The substitution of Gly169 in *B. amyloliquefaciens* subtilisin with Ala and Ser is described in EPO Publication No. 0130756. The same method was used to make the remaining 17 mutants containing all other substituent amino acids for position 169.

The construction protocol is summarized in Figure 18. The overscored and underscored double stranded DNA cassettes used contained the following triplet encoding the substitution of the indicated amino acid at residue 169.

GCT	A	ATG	M
TGT	C	AAC	N
GAT	D	CCT	P
GAA	E	CAA	Q
TTC	F	AGA	R
GGC	G	AGC	S
CAC	H	ACA	T
ATC	I	GTT	V
AAA	K	TGG	W
CTT	L	TAC	Y

Each of the plasmids containing a substituted Gly169 was designated pX169, where X represents the substituent amino acid. The mutant subtilisins were similarly designated.

Two of the above mutant subtilisins, A169 and S169, were analyzed for substrate specificity against synthetic substrates containing Phe, Leu, Ala and Arg in the P-1 position. The following results are shown in Table X.

TABLE X

Effect of Serine and Alanine Mutations at Position 169 on P-1 Substrate Specificity				
Position 169	P-1 Substrate [kcat/Km x 10 ⁻⁴]			
	Phe	Leu	Ala	Arg
Gly (wild type)	40	10	1	0.4
A169	120	20	1	0.9
S169	50	10	1	0.6

These results indicate that substitutions of Ala and Ser at Gly169 have remarkably similar catalytic efficiencies against a range of P-1 substrates compared to their position 166 counterparts. This is probably because position 169 is at the bottom of the P-1 specificity subsite.

EXAMPLE 6

Substitution at Position 104

Tyr104 has been substituted with Ala, His, Leu, Met and Ser. The method used was a modification of the site directed mutagenesis method. According to the protocol of Figure 19, a primer (shaded in line 4) introduced a unique HindIII site and a frame shift mutation at codon 104. Restriction-purification for the unique HindIII site facilitated the isolation of the mutant sequence (line 4). Restriction-selection against this HindIII site using primers in line 5 was used to obtain position 104 mutants.

The following triplets were used in the primers of Figure 19, line 5 for the 104 codon which substituted the following amino acids.

GCT	A	TTC	F
ATG	M	CCT	P
CTT	L	ACA	T
AGC	S	TGG	W
CAC	H	TAC	Y
CAA	Q	GTT	V
GAA	E	AGA	R
GGC	G	AAC	N
ATC	I	GAT	D
AAA	K	TGT	C

EP 0 251 446 B1

The substrates in Table XI were used to analyze the substrate specificity of these mutants. The results obtained for H104 subtilisin are shown in Table XI.

TABLE XI

Substrate	kcat		Km		Kcat/Km	
	WT	H104	WT	H104	WT	H104
sAAPFpNA	50.0	22.0	1.4×10^{-4}	7.1×10^{-4}	3.6×10^5	3.1×10^4
sAAPApNA	3.2	2.0	2.3×10^{-4}	1.9×10^{-3}	1.4×10^4	1×10^3
sFAPFpNA	26.0	38.0	1.8×10^{-4}	4.1×10^{-4}	1.5×10^5	9.1×10^4
sFAPApNA	0.32	2.4	7.3×10^{-5}	1.5×10^{-4}	4.4×10^3	1.6×10^4

From these data it is clear that the substitution of His for Tyr at position 104 produces an enzyme which is more efficient (higher kcat/Km) when Phe is at the P-4 substrate position than when Ala is at the P-4 substrate position.

EXAMPLE 7

Substitution of Ala152

Ala152 has been substituted by Gly and Ser to determine the effect of such substitutions on substrate specificity.

The wild type DNA sequence was mutated by the V152/P153 primer (Figure 20, line 4) using the above restriction-purification approach for the new KpnI site. Other mutant primers (shaded sequences Figure 20; S152, line 5 and G152, line 6) mutated the new KpnI site away and such mutants were isolated using the restriction-selection procedure as described above for loss of the KpnI site.

The results of these substitutions for the above synthetic substrates containing the P-1 amino acids Phe, Leu and Ala are shown in Table XII.

TABLE XII

Position 152	P-1 Substrate (kcat/Km $\times 10^{-4}$)		
	Phe	Leu	Ala
Gly (G)	0.2	0.4	<0.04
Ala (wild type)	40.0	10.0	1.0
Ser (S)	1.0	0.5	0.2

These results indicate that, in contrast to positions 166 and 169, replacement of Ala152 with Ser or Gly causes a dramatic reduction in catalytic efficiencies across all substrates tested. This suggests Ala152, at the top of the S-1 subsite, may be the optimal amino acid because Ser and Gly are homologous Ala substitutes.

EXAMPLE 8

Substitution at Position 156

Mutants containing the substitution of Ser and Gln for Glu156 have been constructed according to the overall method depicted in Figure 21. This method was designed to facilitate the construction of multiple mutants at position 156 and 166 as will be described hereinafter. However, by regenerating the wild type Gly166, single mutations at Glu156 were obtained.

The plasmid p Δ 166 is already depicted in line 2 of Figure 13. The synthetic oligonucleotides at the top right of Figure 21 represent the same DNA cassettes depicted in line 4 of Figure 13. The plasmid p166 in Figure 21 thus represents the mutant plasmids of Examples 3 and 4. In this particular example, p166 contains the wild type Gly166.

Construction of position 156 single mutants were prepared by ligation of the three fragments (1-3) indicated at the bottom of Figure 21. Fragment 3, containing the carboxy-terminal portion of the subtilisin gene including the wild type position 166 codon, was isolated as a 610 bp SacI-BamHI fragment. Fragment 1 contained the vector sequences, as well as the amino-terminal sequences of the subtilisin gene through codon 151. To produce fragment 1, a unique KpnI site at codon 152 was introduced into the wild type subtilisin sequence from pS4.5. Site-directed mutagenesis in M13 employed a primer having the sequence 5'-TA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3' to produce the mutation. Enrichment for the mutant sequence was accomplished by restriction with KpnI, purification and self ligation. The mutant sequence containing the KpnI site was confirmed by direct plasmid sequencing to give pV152. pV152 (~1 µg) was digested with KpnI and treated with 2 units of DNA polymerase I large fragment (Klenow fragment from Boehringer-Mannheim) plus 50 µM deoxynucleotide triphosphates at 37°C for 30 min. This created a blunt end that terminated with codon 151. The DNA was extracted with 1:1 volumes phenol and CHCl₃ and DNA in the aqueous phase was precipitated by addition of 0.1 volumes 5M ammonium acetate and two volumes ethanol. After centrifugation and washing the DNA pellet with 70% ethanol, the DNA was lyophilized. DNA was digested with BamHI and the 4.6kb piece (fragment 1) was purified by acrylamide gel electrophoresis followed by electroelution. Fragment 2 was a duplex synthetic DNA cassette which when ligated with fragments 1 and 3 properly restored the coding sequence except at codon 156. The top strand was synthesized to contain a glutamine codon, and the complementary bottom strand coded for serine at 156. Ligation of heterophosphorylated cassettes leads to a large and favorable bias for the phosphorylated over the non-phosphorylated oligonucleotide sequence in the final segregated plasmid product. Therefore, to obtain Q156 the top strand was phosphorylated, and annealed to the non-phosphorylated bottom strand prior to ligation. Similarly, to obtain S156 the bottom strand was phosphorylated and annealed to the non-phosphorylated top strand. Mutant sequences were isolated after ligation and transformation, and were confirmed by restriction analysis and DNA sequencing as before. To express variant subtilisins, plasmids were transformed into a subtilisin-neutral protease deletion mutant of *B. subtilis*, BG2036, as previously described. Cultures were fermented in shake flasks for 24 h at 37°C in LB media containing 12.5 mg/mL chloramphenicol and subtilisin was purified from culture supernatants as described. Purity of subtilisin was greater than 95% as judged by SDS PAGE.

These mutant plasmids designated pS156 and pQ156 and mutant subtilisins designated S156 and Q156 were analyzed with the above synthetic substrates where P-1 comprised the amino acids Glu, Gln, Met and Lys. The results of this analyses are presented in Example 9.

EXAMPLE 9

Multiple Mutants With Altered Substrate Specificity - Substitution at Positions 156 and 166

Single substitutions of position 166 are described in Examples 3 and 4. Example 8 describes single substitutions at position 156 as well as the protocol of Figure 21 whereby various double mutants comprising the substitution of various amino acids at positions 156 and 166 can be made. This example describes the construction and substrate specificity of subtilisin containing substitutions at position 156 and 166 and summarizes some of the data for single and double mutants at positions 156 and 166 with various substrates.

K166 is a common replacement amino acid in the 156/166 mutants described herein. The replacement of Lys for Gly166 was achieved by using the synthetic DNA cassette at the top right of Figure 21 which contained the triplet AAA for NNN. This produced fragment 2 with Lys substituting for Gly166.

The 156 substituents were Gln and Ser. The Gln and Ser substitutions at Gly156 are contained within fragment 3 (bottom right Figure 21).

The multiple mutants were produced by combining fragments 1, 2 and 3 as described in Example 8. The mutants Q156/K166 and S156/K166 were selectively generated by differential phosphorylation as described. Alternatively, the double 156/166 mutants, c.f. Q156/K166 and S156/K166, were prepared by ligation of the 4.6kb SacI-BamHI fragment from the relevant p156 plasmid containing the 0.6kb SacI-BamHI fragment from the relevant p166 plasmid.

These mutants, the single mutant K166, and the S156 and Q156 mutants of Example 8 were analyzed for substrate specificity against synthetic polypeptides containing Phe or Glu as the P-1 substrate residue. The results are presented in Table XIII.

TABLE XIII

Enzymes Compared (b)	Substrate P-1 Residue	Km		kcat/Km (mutant)	
		kcat	Km	kcat/Km	kcat/Km (wt.)
Glul56/Gly166 (WT)	Phe	50.00	1.4×10^{-4}	3.6×10^5	(1)
	Glu	0.54	3.4×10^{-2}	1.6×10^1	(1)
	Phe	20.00	4.0×10^{-5}	5.2×10^5	1.4
	Glu	0.70	5.6×10^{-5}	1.2×10^4	750
Q156/R166	Phe	30.00	1.9×10^{-5}	1.6×10^6	4.4
	Glu	1.60	3.1×10^{-5}	5.0×10^4	3100
S156/R166	Phe	30.00	1.8×10^{-5}	1.6×10^6	4.4
	Glu	0.60	3.9×10^{-5}	1.6×10^4	1000
S156	Phe	34.00	4.7×10^{-5}	7.3×10^5	2.0
	Glu	0.40	1.8×10^{-3}	1.1×10^2	6.9
E156	Phe	48.00	4.5×10^{-5}	1.1×10^6	3.1
	Glu	0.90	3.3×10^{-3}	2.7×10^2	17

As can be seen in Table XIV, either of these single mutations improve enzyme performance upon substrates with glutamate at the P-1 enzyme binding site. When these single mutations were combined, the resulting multiple enzyme mutants are better than either parent. These single or multiple mutations also alter the relative pH activity profiles of the enzymes as shown in Figure 23.

To isolate the contribution of electrostatics to substrate specificity from other chemical binding forces, these various single and double mutants were analyzed for their ability to bind and cleave synthetic substrates containing Glu, Gln, Met and Lys as the P-1 substrate amino acid. This permitted comparisons between side-chains that were more sterically similar but differed in charge (e.g., Glu versus Gln, Lys versus Met). Similarly, mutant enzymes were assayed against homologous P-1 substrates that were most sterically similar but differed in charge (Table XIV).

TABLE XIV

Kinetics of Position 156/166 Subtilisins
Determined for Different P1 Substrates

Enzyme Position	Net Charge (b)	P-1 Substrate log kcat/Km (log 1/Km) (c)			
		Glu	Gln	Met	Lys
156 166					
Glu Asp	-2	n.d.	3.02 (2.56)	3.93 (2.74)	4.23 (3.00)
Glu Glu	-2	n.d.	3.06 (2.91)	3.86 (3.28)	4.48 (3.69)
Glu Asn	-1	1.62 (2.22)	3.85 (3.14)	4.99 (3.85)	4.15 (2.88)
Glu Gln	-1	1.20 (2.12)	4.36 (3.64)	5.43 (4.36)	4.10 (3.15)
Gln Asp	-1	1.30 (1.79)	3.40 (3.08)	4.94 (3.87)	4.41 (3.22)
Ser Asp	-1	1.23 (2.13)	3.41 (3.09)	4.67 (3.68)	4.24 (3.07)
Glu Met	-1	1.20 (2.30)	3.89 (3.19)	5.64 (4.83)	4.70 (3.89)
Glu Ala	-1	n.d.	4.34 (3.55)	5.65 (4.46)	4.90 (3.24)
Glu Gly (wt)	-1	1.20 (1.47)	3.85 (3.35)	5.07 (3.97)	4.60 (3.13)
Gln Gly	0	2.42 (2.48)	4.53 (3.81)	5.77 (4.61)	3.76 (2.82)
Ser Gly	0	2.31 (2.73)	4.09 (3.68)	5.61 (4.55)	3.46 (2.74)
Gln Asn	0	2.04 (2.72)	4.51 (3.76)	5.79 (4.66)	3.75 (2.74)
Ser Asn	0	1.91 (2.78)	4.57 (3.82)	5.72 (4.64)	3.68 (2.80)
Glu Arg	0	2.91 (3.30)	4.26 (3.50)	5.32 (4.22)	3.19 (2.80)
Glu Lys	0	4.09 (4.25)	4.70 (3.88)	6.15 (4.45)	4.23 (2.93)
Gln Lys	+1	4.70 (4.50)	4.64 (3.68)	5.97 (4.68)	3.23 (2.75)
Ser Lys	+1	4.21 (4.40)	4.84 (3.94)	6.16 (4.90)	3.73 (2.84)

Maximum difference:

log kcat/Km (log 1/Km) (d) 3.5 (3.0) 1.8 (1.4) 2.3 (2.2) -1.3 (-1.0)

Footnotes to Table XIV:

(a) *B. subtilis*, BG 2036, expressing indicated variant subtilisin were fermented and enzymes purified as previously described (Estell, et al. (1985) *J. Biol. Chem.* 260, 6518-6521). Wild type subtilisin is indicated (wt) containing Glu156 and Gly166.

(b) Net charge in the P-1 binding site is defined as the sum of charges from positions 156 and 166 at pH 8.6.

(c) Values for $k_{cat}(s^{-1})$ and $K_m(M)$ were measured in 0.1M Tris pH 8.6 at 25°C as previously described against P-1 substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide, where X is the indicated P-1 amino acid. Values for $\log 1/K_m$ are shown inside parentheses. All errors in determination of k_{cat}/K_m and $1/K_m$ are below 5%.

(d) Because values for Glu156/Asp166(D166) are too small to determine accurately, the maximum difference taken for GluP-1 substrate is limited to a charge range of +1 to -1 charge change.

n.d. = not determined

The k_{cat}/K_m ratios shown are the second order rate constants for the conversion of substrate to product, and represent the catalytic efficiency of the enzyme. These ratios are presented in logarithmic form to scale the data, and because $\log k_{cat}/K_m$ is proportional to the lowering of transition-state activation energy (ΔG^\ddagger). Mutations at position 156 and 166 produce changes in catalytic efficiency toward Glu, Gln, Met and Lys P-1 substrates of 3100, 60, 200 and 20 fold, respectively. Making the P-1 binding-site more positively charged [e.g., compare Gln156/Lys166 (Q156/K166) versus Glu156/Met166 (Glu156/M166)] dramatically increased k_{cat}/K_m toward the Glu P-1 substrate (up to 3100 fold), and decreased the catalytic efficiency toward the Lys P-1 substrate (up to 10 fold). In addition, the results show that the catalytic efficiency of wild type enzyme can be greatly improved toward any of the four P-1 substrates by mutagenesis of the P-1 binding site.

The changes in k_{cat}/K_m are caused predominantly by changes in $1/K_m$. Because $1/K_m$ is approximately equal to $1/K_s$, the enzyme-substrate association constant, the mutations primarily cause a change in substrate binding. These mutations produce smaller effects on k_{cat} that run parallel to the effects on $1/K_m$. The changes in k_{cat} suggest either an alteration in binding in the P-1 binding site in going from the Michaelis-complex E•S to the transition-state complex (E•S[‡]) as previously proposed (Robertus, J.D., et al. (1972) *Biochemistry* 11, 2439-2449; Robertus, J.D., et al. (1972) *Biochemistry* 11, 4293-4303), or change in the position of the scissile peptide bond over the catalytic serine in the E•S complex.

Changes in substrate preference that arise from changes in the net charge in the P-1 binding site show trends that are best accounted for by electrostatic effects (Figure 28). As the P-1 binding cleft becomes more positively charged, the average catalytic efficiency increases much more for the Glu P-1 substrate than for its neutral and isosteric P-1 homolog, Gln (Figure 28A). Furthermore, at the positive extreme both substrates have nearly identical catalytic efficiencies.

In contrast, as the P-1 site becomes more positively charged the catalytic efficiency toward the Lys P-1 substrate decreases, and diverges sharply from its neutral and isosteric homolog, Met (Figure 28B). The similar and parallel upward trend seen with increasing positive charge for the Met and Glu P-1 substrates probably results from the fact that all the substrates are succinylated on their amino-terminal end, and thus carry a formal negative charge.

The trends observed in $\log k_{cat}/K_m$ are dominated by changes in the K_m term (Figures 28C and 28D). As the pocket becomes more positively charged, the $\log 1/K_m$ values converge for Glu and Gln P-1 substrates (Figure 28C), and diverge for Lys and Met P-1 substrates (Figure 28D). Although less

pronounced effects are seen in log k_{cat} , the effects of P-1 charge on log k_{cat} parallel those seen in log $1/K_m$ and become larger as the P-1 pocket becomes more positively charged. This may result from the fact that the transition-state is a tetrahedral anion, and a net positive charge in the enzyme may serve to provide some added stabilization to the transition-state.

5 The effect of the change in P-1 binding-site charge on substrate preference can be estimated from the differences in slopes between the charged and neutral isosteric P-1 substrates (Figure 28B). The average change in substrate preference ($\Delta \log k_{cat}/K_m$) between charged and neutral isosteric substrates increases roughly 10-fold as the complementary charge of the enzyme increases (Table XV). When comparing Glu versus Lys, this difference is 100-fold and the change in substrate preference appears predominantly in the
10 K_m term.

TABLE XV

15	Differential Effect on Binding Site Charge on log k_{cat}/K_m or (log $1/K_m$) for P-1 Substrates that Differ in Charge ^(a)			
	Change in P-1 Binding Site Charge ^(b)	$\Delta \log k_{cat}/K_m$ ($\Delta \log 1/K_m$)		
		GluGln	MetLys	GluLys
20	-2 to -1	n.d.	1.2 (1.2)	n.d.
	-1 to 0	0.7 (0.6)	1.3 (0.8)	2.1 (1.4)
	0 to +1	1.5 (1.3)	0.5 (0.3)	2.0 (1.5)
	Avg. change in log k_{cat}/K_m or (log $1/K_m$) per unit charge change	1.1 (1.0)	1.0 (0.8)	2.1 (1.5)

25 ^(a) The difference in the slopes of curves were taken between the P-1 substrates over the charge interval given for log (k_{cat}/K_m) (Figure 28A, B) and (log $1/K_m$) (Figure 28C, D). Values represent the differential effect a charge change has in distinguishing the substrates that are compared.

^(b) Charge in P-1 binding site is defined as the sum of charges from positions 156 and 166.

30 The free energy of electrostatic interactions in the structure and energetics of salt-bridge formation depends on the distance between the charges and the microscopic dielectric of the media. To dissect these structural and microenvironmental effects, the energies involved in specific salt-bridges were evaluated. In addition to the possible salt-bridges shown (Figures 29A and 29B), reasonable salt-bridges can be built between a Lys P-1 substrate and Asp at position 166, and between a Glu P-1 substrate and a Lys at
35 position 166 (not shown). Although only one of these structures is confirmed by X-ray crystallography (Poulos, T.L., et al. (1976) *J. Mol. Biol.* 257 1097-1103), all models have favorable torsion angles (Sielecki, A.R., et al. (1979) *J. Mol. Biol.* 134, 781-804), and do not introduce unfavorable van der Waals contacts.

The change in charged P-1 substrate preference brought about by formation of the model salt-bridges above are shown in Table XVI.

TABLE XVI

Effect of Salt Bridge Formation Between Enzyme
and Substrate on P1 Substrate Preference (a)

Enzymes Compared (b)		Enzyme Position Changed	P-1 Substrates Compared	Substrate Preference $\Delta \log$ (kcat/Km)		Change in Substrate Preference $\Delta \Delta \log$ (kcat/Km)
1	2			1	2	
Glu156/Asp166	Gln156/Asp166	156	LysMet	+0.30	-0.53	0.83
Glu156/Asn166	Gln156/Asn166	156	LysMet	-0.84	-2.04	1.20
Glu156/Gly166	Gln156/Gly166	156	LysMet	-0.47	-2.10	1.63
Glu156/Lys-166	Gln156/Lys166	156	LysMet	-1.92	-2.74	0.82
				Ave $\Delta \Delta \log$ (kcat/Km)		1.10 \pm 0.3
Glu156/Asp166	Glu156/Asn166	166	LysMet	+0.30	-0.84	1.14
Glu156/Glu166	Glu156/Glu166	166	LysMet	+0.62	-1.33	1.95
Gln156/Asp166	Gln156/Asn166	166	LysMet	-0.53	-2.04	1.51
Ser156/Asp166	Ser156/Asn166	166	LysMet	-0.43	-2.04	1.61
Glu156/Lys166	Glu156/Met166	166	GluGln	-0.63	-2.69	2/06
				Ave $\Delta \Delta \log$ (kcat/Km)		1.70 \pm 0.3

Footnotes to Table XVI:

(a) Molecular modeling shows it is possible to form a salt bridge between the indicated charged P-1 substrate and a complementary charge in the P-1 binding site of the enzyme at the indicated position changed.

(b) Enzymes compared have sterically similar amino acid substitutions that differ in charge at the indicated position.

(c) The P-1 substrates compared are structurally similar but differ in charge. The charged P-1 substrate is complementary to the charge change at the position indicated between enzymes 1 and 2.

(d) Data from Table XIV was used to compute the difference in $\log(k_{cat}/K_m)$ between the charged and the non-charged P-1 substrate (i.e., the substrate preference). The substrate preference is shown separately for enzyme 1 and 2.

(e) The difference in substrate preference between enzyme 1 (more highly charged) and enzyme 2 (more neutral) represents the rate change accompanying the electrostatic interaction.

The difference between catalytic efficiencies (i.e., $\Delta \log k_{cat}/K_m$) for the charged and neutral P-1 substrates (e.g., Lys minus Met or Glu minus Gln) give the substrate preference for each enzyme. The change in substrate preference ($\Delta \Delta \log k_{cat}/K_m$) between the charged and more neutral enzyme homologs (e.g., Glu156/Gly166 minus Gln156(Q156)/Gly166) reflects the change in catalytic efficiency that may be attributed solely to electrostatic effects.

These results show that the average change in substrate preference is considerably greater when electrostatic substitutions are produced at position 166 (50-fold in k_{cat}/K_m) versus position 156 (12-fold in k_{cat}/K_m). From these $\Delta \Delta \log k_{cat}/K_m$ values, an average change in transition-state stabilization energy can be calculated of -1.5 and -2.4 kcal/mol for substitutions at positions 156 and 166, respectively. This should represent the stabilization energy contributed from a favorable electrostatic interaction for the binding of free enzyme and substrate to form the transition-state complex.

EXAMPLE 10Substitutions at Position 217

Tyr217 has been substituted by all other 19 amino acids. Cassette mutagenesis as described in EPO publication No. 0130756 was used according to the protocol of Figure 22. The EcoRV restriction site was used for restriction-purification of p Δ 217.

Since this position is involved in substrate binding, mutations here effect kinetic parameters of the enzyme. An example is the substitution of Leu for Tyr at position 217. For the substrate sAAPFPNa, this mutant has a k_{cat} of 277 s^{-1} and a K_m of 4.7×10^{-4} with a k_{cat}/K_m ratio of 6×10^5 . This represents a 5.5-fold increase in k_{cat} with a 3-fold increase in K_m over the wild type enzyme.

In addition, replacement of Tyr217 by Lys, Arg, Phe or Leu results in mutant enzymes which are more stable at pHs of about 9-11 than the WT enzyme. Conversely, replacement of Tyr217 by Asp, Glu, Gly or Pro results in enzymes which are less stable at pHs of about 9-11 than the WT enzyme.

EXAMPLE 11

Multiple Mutants Having Altered Thermal Stability

- 5 B. amyloliquefacien subtilisin does not contain any cysteine residues. Thus, any attempt to produce thermal stability by Cys cross-linkage required the substitution of more than one amino acid in subtilisin with Cys. The following subtilisin residues were multiply substituted with cysteine:

Thr22/Ser87

Ser24/Ser87

- 10 Mutagenesis of Ser24 to Cys was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

15 5'-pC-TAC-ACT-GGA-TGC-AAT-GTT-AAA-G-3'.

- (Asterisks show the location of mismatches and the underlined sequence shows the position of the altered Sau3A site.) The B. amyloliquefaciens subtilisin gene on a 1.5 kb EcoRI-BAMHI fragment from pS4.5 was cloned into M13mp11 and single stranded DNA was isolated. This template (M13mp11SUBT) was double primed with the 5' phosphorylated M13 universal sequencing primer and the mutagenesis primer. Adelman, et al. (1983) DNA 2, 183-193. The heteroduplex was transfected into competent JM101 cells and plaques were probed for the mutant sequence (Zoller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500; Wallace, et al. (1981) Nucleic Acid Res. 9, 3647-3656) using a tetramethylammonium chloride hybridization protocol (Wood, et al. (1985) Proc. Natl. Acad. Sci. USA 82, 1585-1588). The Ser87 to Cys mutation was prepared in a similar fashion using a 5' phosphorylated primer having the sequence

30 5'-pGGC-GTT-GCG-CCA-TGC-GCA-TCA-CT-3'.

- (The asterisk indicates the position of the mismatch and the underlined sequence shows the position of a new MstI site.) The C24 and C87 mutations were obtained at a frequency of one and two percent, respectively. Mutant sequences were confirmed by dideoxy sequencing in M13.

- 35 Mutagenesis of Tyr21/Thr22 to A21/C22 was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

40 5'-pAC-TCT-CAA-GGC-GCT-TGT-GGC-TCA-AAT-GTT-3'.

- (The asterisks show mismatches to the wild type sequence and the underlined sequence shows the position of an altered Sau3A site.) Manipulations for heteroduplex synthesis were identical to those described for C24. Because direct cloning of the heteroduplex DNA fragment can yield increased frequencies of mutagenesis, the EcoRI-BamHI subtilisin fragment was purified and ligated into pBS42. E. coli MM 294 cells were transformed with the ligation mixture and plasmid DNA was purified from isolated transformants. Plasmid DNA was screened for the loss of the Sau3A site at codon 23 that was eliminated by the mutagenesis primer. Two out of 16 plasmid preparations had lost the wild type Sau3A site. The mutant sequence was confirmed by dideoxy sequencing in M13.

- Double mutants, C22/C87 and C24/C87, were constructed by ligating fragments sharing a common Clal site that separated the single parent cysteine codons. Specifically, the 500 bp EcoRI-Clal fragment containing the 5' portion of the subtilisin gene (including codons 22 and 24) was ligated with the 4.7 kb Clal-EcoRI fragment that contained the 3' portion of the subtilisin gene (including codon 87) plus pBS42 vector sequence. E. coli MM 294 was transformed with ligation mixtures and plasmid DNA was purified from individual transformants. Double-cysteine plasmid constructions were identified by restriction site markers originating from the parent cysteine mutants (i.e., C22 and C24, Sau3A minus; Cys87, MstI plus). Plasmids from E. coli were transformed into B. subtilis BG2036. The thermal stability of these mutants as compared to wild type subtilisin are presented in Figure 30 and Tables XVII and XVIII.

TABLE XVII

Effect of DTT on the Half-Time of Autolytic Inactivation of Wild-Type and Disulfide Mutants of Subtilisin*			
Enzyme	t _{1/2}		-DTT/ + DTT
	-DDT	+ DTT	
	min		
Wild-type	95	85	1.1
C22/C87	44	25	1.8
C24/C87	92	62	1.5

(*) Purified enzymes were either treated or not treated with 25mM DTT and dialyzed with or without 10mM DTT in 2mM CaCl₂, 50mM Tris (pH 7.5) for 14 hr. at 4 ° C. Enzyme concentrations were adjusted to 80μl aliquots were quenched on ice and assayed for residual activity. Half-times for autolytic inactivation were determined from semi-log plots of log₁₀ (residual activity) versus time. These plots were linear for over 90% of the inactivation.

TABLE XVIII

Effect of Mutations in Subtilisin on the Half-Time of Autolytic Inactivation at 58 ° C*	
Enzyme	$t_{1/2}$
	min
Wild-type	120
C22	22
C24	120
C87	104
C22/C87	43
C24/C87	115

(*) Half-times for autolytic inactivation were determined for wild-type and mutant subtilisins as described in the legend to Table III. Unpurified and non-reduced enzymes were used directly from *B. subtilis* culture supernatants.

The disulfides introduced into subtilisin did not improve the autolytic stability of the mutant enzymes when compared to the wild-type enzyme. However, the disulfide bonds did provide a margin of autolytic stability when compared to their corresponding reduced double-cysteine enzyme. Inspection of a highly refined x-ray structure of wild-type *B. amyloliquefaciens* subtilisin reveals a hydrogen bond between Thr22 and Ser87. Because cysteine is a poor hydrogen donor or acceptor (Paul, I.C. (1974) in *Chemistry of the -SH Group* (Patai, S., ed.) pp. 111-149, Wiley Interscience, New York) weakening of 22/87 hydrogen bond may explain why the C22 and C87 single-cysteine mutant proteins are less autolytically stable than either C24 or wild-type (Table XVIII). The fact that C22 is less autolytically stable than C87 may be the result of the Tyr21A mutation (Table XVIII). Indeed, construction and analysis of Tyr21/C22 shows the mutant protein has an autolytic stability closer to that of C87. In summary, the C22 and C87 of single-cysteine mutations destabilize the protein toward autolysis, and disulfide bond formation increases the stability to a level less than or equal to that of wild-type enzyme.

EXAMPLE 12

Multiple Mutants Containing Substitutions at Position 222 and Position 166 or 169

Double mutants 166/222 and 169/222 were prepared by ligating together (1) the 2.3kb *A*cclI fragment from pS4.5 which contains the 5' portion of the subtilisin gene and vector sequences, (2) the 200bp *A*vaII fragment which contains the relevant 166 or 169 mutations from the respective 166 or 169 plasmids, and (3) the 2.2kb *A*vaII fragment which contains the relevant 222 mutation 3' and of the subtilisin genes and vector

sequence from the respective p222 plasmid.

Although mutations at position 222 improve oxidation stability they also tend to increase the K_m . An example is shown in Table XIX. In this case the A222 mutation was combined with the K166 mutation to give an enzyme with k_{cat} and K_m intermediate between the two parent enzymes.

TABLE XIX

	k_{cat}	K_m
WT	50	1.4×10^{-4}
A222	42	9.9×10^{-4}
K166	21	3.7×10^{-5}
K166/A222	29	2.0×10^{-4}
substrate sAAPFPNa		

EXAMPLE 13

Multiple Mutants Containing Substitutions at Positions 50, 156, 166, 217 and Combinations Thereof

The double mutant S156/A169 was prepared by ligation of two fragments, each containing one of the relevant mutations. The plasmid pS156 was cut with XmaI and treated with S1 nuclease to create a blunt end at codon 167. After removal of the nuclease by phenol/chloroform extraction and ethanol precipitation, the DNA was digested with BamHI and the approximately 4kb fragment containing the vector plus the 5' portion of the subtilisin gene through codon 167 was purified.

The pA169 plasmid was digested with KpnI and treated with DNA polymerase Klenow fragment plus 50 μ M dNTPs to create a blunt end codon at codon 168. The Klenow was removed by phenol/chloroform extraction and ethanol precipitation. The DNA was digested with BamHI and the 590bp fragment including codon 168 through the carboxy terminus of the subtilisin gene was isolated. The two fragments were then ligated to give S156/A169.

Triple and quadruple mutants were prepared by ligating together (1) the 220bp PvuII/HaeIII fragment containing the relevant 156, 166 and/or 169 mutations from the respective p156, p166 and/or p169 double of single mutant plasmid, (2) the 550bp HaeIII/BamHI fragment containing the relevant 217 mutant from the respective p217 plasmid, and (3) the 3.9kb PvuII/BamHI fragment containing the F50 mutation and vector sequences.

The multiple mutant F50/S156/A169/L217, as well as B. amyloliquefaciens subtilisin, B. licheniformis subtilisin and the single mutant L217 were analyzed with the above synthetic polypeptides where the P-1 amino acid in the substrate was Lys, His, Ala, Gln, Tyr, Phe, Met and Leu. These results are shown in Figures 26 and 27.

These results show that the F50/S156/A169/L217 mutant has substrate specificity similar to that of the B. licheniformis enzyme and differs dramatically from the wild type enzyme. Although only data for the L217 mutant are shown, none of the single mutants (e.g., F50, S156 or A169) showed this effect. Although B. licheniformis differs in 88 residue positions from B. amyloliquefaciens, the combination of only these four mutations accounts for most of the differences in substrate specificity between the two enzymes.

EXAMPLE 14

Subtilisin Mutants Having Altered Alkaline Stability

A random mutagenesis technique was used to generate single and multiple mutations within the B. amyloliquefaciens subtilisin gene. Such mutants were screened for altered alkaline stability. Clones having increased (positive) alkaline stability and decreased (negative) alkaline stability were isolated and sequenced to identify the mutations within the subtilisin gene. Among the positive clones, the mutants V107 and R213 were identified. These single mutants were subsequently combined to produce the mutant V107/R213.

One of the negative clones (V50) from the random mutagenesis experiments resulted in a marked decrease in alkaline stability. Another mutant (P50) was analyzed for alkaline stability to determine the effect

of a different substitution at position 50. The F50 mutant was found to have a greater alkaline stability than wild type subtilisin and when combined with the double mutant V107/R213 resulted in a mutant having an alkaline stability which reflected the aggregate of the alkaline stabilities for each of the individual mutants.

The single mutant R204 and double mutant C204/R213 were identified by alkaline screening after random cassette mutagenesis over the region from position 197 to 228. The C204/R213 mutant was thereafter modified to produce mutants containing the individual mutations C204 and R213 to determine the contribution of each of the individual mutations. Cassette mutagenesis using pooled oligonucleotides to substitute all amino acids at position 204, was utilized to determine which substitution at position 204 would maximize the increase in alkaline stability. The mutation from Lys213 to Arg was maintained constant for each of these substitutions at position 204.

A. Construction of pB0180, an *E. coli*-*B. subtilis* Shuttle Plasmid

The 2.9 kb EcoRI-BamHI fragment from pBR327 (Covarrubias, L., et al. (1981) *Gene* **13**, 25-35) was ligated to the 3.7kb EcoRI-BamHI fragment of pBD64 (Gryczan, T., et al. (1980) *J. Bacteriol.*, **141**, 246-253) to give the recombinant plasmid pB0153. The unique EcoRI recognition sequence in pBD64 was eliminated by digestion with EcoRI followed by treatment with Klenow and deoxynucleotide triphosphates (Maniatis, T., et al. (eds.) (1982) in *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). Blunt end ligation and transformation yielded pB0154. The unique Aval recognition sequence in pB0154 was eliminated in a similar manner to yield pB0171. pB0171 was digested with BamHI and PvuII and treated with Klenow and deoxynucleotide triphosphates to create blunt ends. The 6.4 kb fragment was purified, ligated and transformed into LE392 cells (Enquist, L.W., et al. (1977) *J. Mol. Biol.* **111**, 97-120), to yield pB0172 which retains the unique BamHI site. To facilitate subcloning of subtilisin mutants, a unique and silent KpnI site starting at codon 166 was introduced into the subtilisin gene from pS4.5 (Wells, J.A., et al. (1983) *Nucleic Acids Res.*, **11**, 7911-7925) by site-directed mutagenesis. The KpnI + plasmid was digested with EcoRI and treated with Klenow and deoxynucleotide triphosphates to create a blunt end. The Klenow was inactivated by heating for 20 min at 68 °C, and the DNA was digested with BamHI. The 1.5 kb blunt EcoRI-BamHI fragment containing the entire subtilisin was ligated with the 5.8 kb NruI-BamHI from pB0172 to yield pB0180. The ligation of the blunt NruI end to the blunt EcoRI end recreated an EcoRI site. Proceeding clockwise around pB0180 from the EcoRI site at the 5' end of the subtilisin gene is the unique BamHI site at the 3' end of the subtilisin gene, the chloramphenicol and neomycin resistance genes and UB110 gram positive replication origin derived from pBD64, the ampicillin resistance gene and gram negative replication origin derived from pBR327.

B. Construction of Random Mutagenesis Library

The 1.5 kb EcoRI-BamHI fragment containing the *B. amyloliquefaciens* subtilisin gene (Wells et al., 1983) from pB0180 was cloned into M13mp11 to give M13mp11 SUBT essentially as previously described (Wells, J.A., et al. (1986) *J. Biol. Chem.*, **261**, 6564-6570). Deoxyuridine containing template DNA was prepared according to Kunkel (Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82** 488-492). Uridine containing template DNA (Kunkel, 1985) was purified by CsCl density gradients (Maniatis, T. et al. (eds.) (1982) in *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). A primer (Aval⁻) having the sequence

5' GAAAAAAGACCCCTAGCGTCGCTTA

ending at codon -11, was used to alter the unique Aval recognition sequence within the subtilisin gene. (The asterisk denotes the mismatches from the wild-type sequence and underlined is the altered Aval site.)

The 5' phosphorylated Aval primer (~320 pmol) and ~40 pmol (~120 µg) of uridine containing M13mp11 SUBT template in 1.88 ml of 53 mM NaCl, 7.4 mM MgCl₂ and 7.4 mM Tris.HCl (pH 7.5) were annealed by heating to 90 °C for 2 min. and cooling 15 min at 24 °C (Fig. 31). Primer extension at 24 °C was initiated by addition of 100 µL containing 1 mM in all four deoxynucleotide triphosphates, and 20 µL Klenow fragment (5 units/µl). The extension reaction was stopped every 15 seconds over ten min by addition of 10 µL 0.25 M EDTA (pH 8) to 50 µL aliquots of the reaction mixture. Samples were pooled, phenol chloroform extracted and DNA was precipitated twice by addition of 2.5 vol 100% ethanol, and washed twice with 70% ethanol.

The pellet was dried, and redissolved in 0.4 ml 1 mM EDTA, 10 mM Tris (pH 8).

Misincorporation of α -thiodeoxynucleotides onto the 3' ends of the pool of randomly terminated template was carried out by incubating four 0.2 ml solutions each containing one-fourth of the randomly terminated template mixture (~20 μ g), 0.25 mM of a given α -thiodeoxynucleotide triphosphate, 100 units AMV polymerase, 50 mM KCL, 10 mM $MgCl_2$, 0.4 mM dithiothreitol, and 50 mM Tris (pH 8.3) (Champoux, J.J. (1984) *Genetics*, 2, 454-464). After incubation at 37°C for 90 minutes, misincorporation reactions were sealed by incubation for five minutes at 37°C with 50 mM all four deoxynucleotide triphosphates (pH 8), and 50 units AMV polymerase. Reactions were stopped by addition of 25 mM EDTA (final), and heated at 68°C for ten min to inactivate AMV polymerase. After ethanol precipitation and resuspension, synthesis of closed circular heteroduplexes was carried out for two days at 14°C under the same conditions used for the timed extension reactions above, except the reactions also contained 1000 units T4 DNA ligase, 0.5 mM ATP and 1 mM β -mercaptoethanol. Simultaneous restriction of each heteroduplex pool with KpnI, BamHI, and EcoRI confirmed that the extension reactions were nearly quantitative. Heteroduplex DNA in each reaction mixture was methylated by incubation with 80 μ M S-adenosylmethionine and 150 units dam methylase for 1 hour at 37°C. Methylation reactions were stopped by heating at 68°C for 15 min.

One-half of each of the four methylated heteroduplex reactions were transformed into 2.5 ml competent *E. coli* JM101 (Messing, J. (1979) *Recombinant DNA Tech. Bull.*, 2, 43-48). The number of independent transformants from each of the four transformations ranged from $0.4-2.0 \times 10^5$. After growing out phage pools, RF DNA from each of the four transformations was isolated and purified by centrifugation through CsCl density gradients. Approximately 2 μ g of RF DNA from each of the four pools was digested with EcoRI, BamHI and Aval. The 1.5 kb EcoRI-BamHI fragment (i.e., Aval resistant) was purified on low gel temperature agarose and ligated into the 5.5 kb EcoRI-BamHI vector fragment of pB0180. The total number of independent transformants from each α -thiodeoxynucleotide misincorporation plasmid library ranged from $1.2-2.4 \times 10^4$. The pool of plasmids from each of the four transformations was grown out in 200 ml LB media containing 12.5 μ g/ml cmp and plasmid DNA was purified by centrifugation through CsCl density gradients.

C. Expression and Screening of Subtilisin Point Mutants

Plasmid DNA from each of the four misincorporation pools was transformed (Anagnostopoulos, C., et al. (1967), *J. Bacteriol.*, 81, 741-746) into BG2036. For each transformation, 5 μ g of DNA produced approximately 2.5×10^5 independent BG2036 transformants, and liquid culture aliquots from the four libraries were stored in 10% glycerol at 70°C. Thawed aliquots of frozen cultures were plated on LB/5 μ g/ml cmp/1.6% skim milk plates (Wells, J.A., et al. (1983) *Nucleic Acids Res.*, 11, 7911-7925), and fresh colonies were arrayed onto 96-well microtiter plates containing 150 l per well LB media plus 12.5 μ g/ml cmp. After 1 h at room temperature, a replica was stamped (using a matched 96 prong stamp) onto a 132 mm BA 85 nitrocellulose filter (Schleicher and Scheull) which was layered on a 140 mm diameter LB/cmp/skim milk plate. Cells were grown about 16 h at 30°C until halos of proteolysis were roughly 5-7 mm in diameter and filters were transferred directly to a freshly prepared agar plate at 37°C containing only 1.6% skim milk and 50 mM sodium phosphate pH 11.5. Filters were incubated on plates for 3-6 h at 37°C to produce halos of about 5 mm for wild-type subtilisin and were discarded. The plates were stained for 10 min at 24°C with Coomassie blue solution (0.25% Coomassie blue (R-250) 25% ethanol) and destained with 25% ethanol, 10% acetic acid for 20 min. Zones of proteolysis appeared as blue halos on a white background on the underside of the plate and were compared to the original growth plate that was similarly stained and destained as a control. Clones were considered positive that produced proportionately larger zones of proteolysis on the high pH plates relative to the original growth plate. Negative clones gave smaller halos under alkaline conditions. Positive and negative clones were restreaked to colony purify and screened again in triplicate to confirm alkaline pH results.

D. Identification and Analysis of Mutant Subtilisins

Plasmid DNA from 5 ml overnight cultures of more alkaline active *B. subtilis* clones was prepared according to Birnboim and Doly (Birnboim, H.C., et al. (1979) *Nucleic Acid Res.* 7, 1513) except that incubation with 2 mg/ml lysozyme proceeded for 5 min at 37°C to ensure cell lysis and an additional phenol/ $CHCl_3$ extraction was employed to remove contaminants. The 1.5 kb EcoRI-BamHI fragment containing the subtilisin gene was ligated into M13mp11 and template DNA was prepared for DNA sequencing (Messing, J., et al. (1982) *Gene*, 19 269-276). Three DNA sequencing primers ending at codon 26, +95, and +155 were synthesized to match the subtilisin coding sequence. For preliminary sequence

identification a single track of DNA sequence, corresponding to the dNTPas misincorporation library from which the mutant came, was applied over the entire mature protein coding sequence (i.e., a single dideoxyguanosine sequence track was applied to identify a mutant from the dGTPas library). A complete four track of DNA sequence was performed 200 bp over the site of mutagenesis to confirm and identify the mutant sequence (Sanger, F., et al., (1980) *J. Mol. Biol.*, **143**, 161-178). Confirmed positive and negative bacilli clones were cultured in LB media containing 12.5µg/mL cmp and purified from culture supernatants as previously described (Estell, D.A., et al. (1985) *J. Biol. Chem.*, **260**, 6518-6521). Enzymes were greater than 98% pure as analyzed by SDS-polyacrylamide gel electrophoresis (Laemmli, U.K. (1970), *Nature*, **227**, 680-685), and protein concentrations were calculated from the absorbance at 280 nm,

10

$$\epsilon_{280}^{0.1\%} = 1.17$$

15 (Maturbara, H., et al. (1965), *J. Biol. Chem.*, **240**, 1125-1130).

Enzyme activity was measured with 200µg/mL succinyl-L-AlaL-AlaL-ProL-Phep-nitroanilide (Sigma) in 0.1M Tris pH 8.6 or 0.1 M CAPS pH 10.8 at 25°C. Specific activity (µ moles product/min-mg) was calculated from the change in absorbance at 410 nm from production of p-nitroaniline with time per mg of enzyme (E410 = 8,480 M-lcm-l; Del Mar, E.G., et al. (1979), *Anal. Biochem.*, **99**, 316-320). Alkaline autolytic stability studies were performed on purified enzymes (200µg/mL) in 0.1 M potassium phosphate (pH 12.0) at 37°C. At various times aliquots were assayed for residual enzyme activity (Wells, J.A., et al. (1986) *J. Biol. Chem.*, **261**, 6564-6570).

E. Results

25

1. Optimization and analysis of mutagenesis frequency

A set of primer-template molecules that were randomly 3'-terminated over the subtilisin gene (Fig. 31) was produced by variable extension from a fixed 5'-primer (The primer mutated a unique Aval site at codon 11 in the subtilisin gene). This was achieved by stopping polymerase reactions with EDTA after various times of extension. The extent and distribution of duplex formation over the 1 kb subtilisin gene fragment was assessed by multiple restriction digestion (not shown). For example, production of new HinfI fragments identified when polymerase extension had proceeded past Ile110, Leu233, and Asp259 in the subtilisin gene.

Misincorporation of each dNTPas at randomly terminated 3' ends by AMV reverse transcriptase (Zakour, R.A., et al. (1982), *Nature*, **295**, 708-710; Zakour, R.A., et al. (1984), *Nucleic Acids Res.*, **12**, 6615-6628) used conditions previously described (Champoux, J.J., (1984), *Genetics*, **2**, 454-464). The efficiency of each misincorporation reaction was estimated to be greater than 80% by the addition of each dNTPas to the Aval restriction primer, and analysis by polyacrylamide gel electrophoresis. Misincorporations were sealed by polymerization with all four dNTP's and closed circular DNA was produced by reaction with DNA ligase.

Several manipulations were employed to maximize the yield of the mutant sequences in the heteroduplex. These included the use of a deoxyuridine containing template (Kunkel, T.A. (1985), *Proc. Natl. Acad. Sci. USA*, **82** 488-492; Pukkila, P.J. et al. (1983), *Genetics*, **104**, 571-582), *in vitro* methylation of the mutagenic strand (Kramer, W. et al. (1982) *Nucleic Acids Res.*, **10** 6475-6485), and the use of Aval restriction-selection against the wild-type template strand which contained a unique Aval site. The separate contribution of each of these enrichment procedures to the final mutagenesis frequency was not determined, except that prior to Aval restriction-selection roughly one-third of the segregated clones in each of the four pools still retained a wild-type Aval site within the subtilisin gene. After Aval restriction-selection greater than 98% of the plasmids lacked the wild-type Aval site.

The 1.5 kb EcoRI-BamHI subtilisin gene fragment that was resistant to Aval restriction digestion, from each of the four CsCl purified M13 RF pools was isolated on low melting agarose. The fragment was ligated *in situ* from the agarose with a similarly cut *E. coli*-B. subtilis shuttle vector, pB0180, and transformed directly into *E. coli* LE392. Such direct ligation and transformation of DNA isolated from agarose avoided losses and allowed large numbers of recombinants to be obtained (>100,000 per µg equivalent of input M13 pool).

The frequency of mutagenesis for each of the four dNTPas misincorporation reactions was estimated from the frequency that unique restriction sites were eliminated (Table XX). The unique restriction sites

chosen for this analysis, Clal, PvuII, and KpnI, were distributed over the subtilisin gene starting at codons 35, 104, and 166, respectively. As a control, the mutagenesis frequency was determined at the PstI site located in the β lactamase gene which was outside the window of mutagenesis. Because the absolute mutagenesis frequency was close to the percentage of undigested plasmid DNA, two rounds of restriction-selection were necessary to reduce the background of surviving uncut wild-type plasmid DNA below the mutant plasmid (Table XX). The background of surviving plasmid from wild-type DNA probably represents the sum total of spontaneous mutations, uncut wild-type plasmid, plus the efficiency with which linear DNA can transform *E. coli*. Subtracting the frequency for unmutagenized DNA (background) from the frequency for mutant DNA, and normalizing for the window of mutagenesis sampled by a given restriction analysis (4-6 bp) provides an estimate of the mutagenesis efficiency over the entire coding sequence (~1000 bp).

15

20

25

30

35

40

45

50

55

TABLE XX

5	α -thiol dNTP misincor- porated (b)	Restriction Site Selection	% resistant clones ^c			% resistant clones over Background ^d	% mutants per 1000bp ^e
			1st round	2nd round	Total		
10	None	<u>PstI</u>	0.32	0.7	0.002	0	-
	G	<u>PstI</u>	0.33	1.0	0.003	0.001	0.2
	T	<u>PstI</u>	0.32	<0.5	<0.002	0	0
	C	<u>PstI</u>	0.43	3.0	0.013	0.011	3
15	None	<u>ClaI</u>	0.28	5	0.014	0	-
	G	<u>ClaI</u>	2.26	85	1.92	1.91	380
	T	<u>ClaI</u>	0.48	31	0.15	0.14	35
	C	<u>ClaI</u>	0.55	15	0.08	0.066	17
20	None	<u>PvuII</u>	0.08	29	0.023	0	-
	G	<u>PvuII</u>	0.41	90	0.37	0.35	88
	T	<u>PvuII</u>	0.10	67	0.067	0.044	9
	C	<u>PvuII</u>	0.76	53	0.40	0.38	95
25	None	<u>KpnI</u>	0.41	3	0.012	0	-
	G	<u>KpnI</u>	0.98	35	0.34	0.33	83
	T	<u>KpnI</u>	0.36	15	0.054	0.042	8
	C	<u>KpnI</u>	1.47	26	0.38	0.37	93
30							
35							

(a) Mutagenesis frequency is estimated from the frequency for obtaining mutations that alter unique restriction sites within the mutagenized subtilisin gene (i.e., ClaI, PvuII, or KpnI) compared to mutation frequencies of the PstI site, that is outside the window of mutagenesis.

(b) Plasmid DNA was from wild-type (none) or mutagenized by dNTPs misincorporation as described.

(c) Percentage of resistant clones was calculated from the fraction of clones obtained after three fold or greater over-digestion of the plasmid with the indicated restriction enzyme compared to a

non-digested control. Restriction-resistant plasmid DNA from the first round was subjected to a second round of restriction-selection. The total represents the product of the fractions of resistant clones obtained from both rounds of selection and gives percentage of restriction-site mutant clones in the original starting pool. Frequencies were derived from counting at least 20 colonies and usually greater than 100.

(d) Percent resistant clones was calculated by subtracting the percentage of restriction-resistant clones obtained for wild-type DNA (i.e., none) from that obtained for mutant DNA.

(e) This extrapolates from the frequency of mutation over each restriction site to the entire subtilisin gene (~1 kb). This has been normalized to the number of possible bases (4-6 bp) within each restriction site that can be mutagenized by a given misincorporation event.

From this analysis, the average percentage of subtilisin genes containing mutations that result from dGTP α s, dCTP α s, or dTTP α s misincorporation was estimated to be 90, 70, and 20 percent, respectively. These high mutagenesis frequencies were generally quite variable depending upon the dNTP α s and misincorporation efficiencies at this site. Misincorporation efficiency has been reported to be both dependent on the kind of mismatch, and the context of primer (Champoux, J.J., (1984); Skinner, J.A., et al. (1986) *Nucleic Acids Res.*, 14, 6945-6964). Biased misincorporation efficiency of dGTP α s and dCTP α s over dTTP α s has been previously observed (Shortle, D., et al. (1985), *Genetics*, 110, 539-555). Unlike the dGTP α s, dCTP α s, and dTTP α s libraries the efficiency of mutagenesis for the dATP α s misincorporation library could not be accurately assessed because 90% of the restriction-resistant plasmids analyzed simply lacked the subtilisin gene insert. This problem probably arose from self-ligation of the vector when the dATP α s mutagenized subtilisin gene was subcloned from M13 into pB0180. Correcting for the vector background, we estimate the mutagenesis frequency around 20 percent in the dATP α s misincorporation library. In a separate experiment (not shown), the mutagenesis efficiencies for dGTP α s and dTTP α s misincorporation were estimated to be around 50 and 30 percent, respectively, based on the frequency of reversion of an inactivating mutation at codon 169.

The location and identity of each mutation was determined by a single track of DNA sequencing corresponding to the misincorporated α thiodeoxynucleotide over the entire gene followed by a complete four track of DNA sequencing focused over the site of mutation. Of 14 mutants identified, the distribution was similar to that reported by Shortle and Lin (1985) except we did not observe nucleotide insertion or deletion mutations. The proportion of AG mutations was highest in the G misincorporation library, and some unexpected point mutations appeared in the dTTP α s and dCTP α s libraries.

2. Screening and Identification of Alkaline Stability Mutants of Subtilisin

It is possible to screen colonies producing subtilisin by halos of casein digestion (Wells, J.A. et al. (1983) *Nucleic Acids Res.*, 11, 7911-7925). However, two problems were posed by screening colonies under high alkaline conditions (>pH 11). First, *B. subtilis* will not grow at high pH, and we have been unable to transform an alkylphilic strain of bacillus. This problem was overcome by adopting a replica plating strategy in which colonies were grown on filters at neutral pH to produce subtilisin and filters subsequently transferred to casein plates at pH 11.5 to assay subtilisin activity. However, at pH 11.5 the casein micells no longer formed a turbid background and thus prevented a clear observation of proteolysis halos. The problem was overcome by briefly staining the plate with Coomassie blue to amplify proteolysis zones and acidifying the plates to develop casein micell turbidity. By comparison of the halo size produced on the reference growth plate (pH 7) to the high pH plate (pH 11.5), it was possible to identify mutant subtilisins that had increased (positives) or decreased (negatives) stability under alkaline conditions.

Roughly 1000 colonies were screened from each of the four misincorporation libraries. The percentage of colonies showing a differential loss of activity at pH 11.5 versus pH 7 represented 1.4, 1.8, 1.4, and 0.6% of the total colonies screened from the thiol dGTP α s, dATP α s, dTTP α s, and dCTP α s libraries, respectively. Several of these negative clones were sequenced and all were found to contain a single base change as expected from the misincorporation library from which they came. Negative mutants included A36, E170 and V50. Two positive mutants were identified as V107 and R213. The ratio of negatives to positives was roughly 50:1.

3. Stability and Activity of Subtilisin Mutants at Alkaline pH

Subtilisin mutants were purified and their autolytic stabilities were measured by the time course of inactivation at pH 12.0 (Figs. 32 and 33). Positive mutants identified from the screen (i.e., V107 and R213) were more resistant to alkaline induced autolytic inactivation compared to wild-type; negative mutants (i.e., E170 and V50) were less resistant. We had advantageously produced another mutant at position 50 (F50) by site-directed mutagenesis. This mutant was more stable than wild-type enzyme to alkaline autolytic inactivation (Fig. 33). At the termination of the autolysis study, SDS-PAGE analysis confirmed that each subtilisin variant had autolyzed to an extent consistent with the remaining enzyme activity.

The stabilizing effects of V107, R213, and F50 are cumulative. See Table XXI. The double mutant, V107/R213 (made by subcloning the 920 bp *EcoRI*-*KpnI* fragment of pB0180V107 into the 6.6 kb *EcoRI*-*KpnI* fragment of pB0180R213), is more stable than either single mutant. The triple mutant, F50/V107/R213 (made by subcloning the 735 bp *EcoRI*-*PvuII* fragment of pF50 (Example 2) into the 6.8 kb *EcoRI*-*PvuII* fragment of pB0180V107), is more stable than the double mutant V107/R213 or F50. The inactivation curves show a biphasic character that becomes more pronounced the more stable the mutant analyzed. This may result from some destabilizing chemical modification(s) (eg., deamidation) during the autolysis study and/or reduced stabilization caused by complete digestion of larger autolysis peptides. These alkaline autolysis studies have been repeated on separately purified enzyme batches with essentially the same results. Rates of autolysis should depend both on the conformational stability as well as the specific activity of the subtilisin variant (Wells, J.A., et al. (1986), *J. Biol. Chem.*, **261**, 6564-6570). It was therefore possible that the decreases in autolytic inactivation rates may result from decreases in specific activity of the more stable mutant under alkaline conditions. In general the opposite appears to be the case. The more stable mutants, if anything, have a relatively higher specific activity than wild-type under alkaline conditions and the less stable mutants have a relatively lower specific activity. These subtle effects on specific activity for V107/R213 and F50/V107/R213 are cumulative at both pH 8.6 and 10.8. The changes in specific activity may reflect slight differences in substrate specificity, however, it is noteworthy that only positions 170 and 107 are within 6Å of a bound model substrate (Robertus, J.D., et al. (1972), *Biochemistry* **11**, 2438-2449).

TABLE XXI

Relationship between relative specific activity at pH 8.6 or 10.8 and alkaline autolytic stability			
Enzyme	Relative specific activity		Alkaline autolysis half-time (min) ^b
	pH 8.6	pH 10.8	
Wild-type	100 \pm 1	100 \pm 3	86
Q170	46 \pm 1	28 \pm 2	13
V107	126 \pm 3	99 \pm 5	102
R213	97 \pm 1	102 \pm 1	115
V107/R213	116 \pm 2	106 \pm 3	130
V50	66 \pm 4	61 \pm 1	58
F50	123 \pm 3	157 \pm 7	131
F50/V107/R213	126 \pm 2	152 \pm 3	168

^(a) Relative specific activity was the average from triplicate activity determinations divided by the wild-type value at the same pH. The average specific activity of wild-type enzyme at pH 8.6 and 10.8 was 70 μ moles/min-mg and 37 μ moles/min-mg, respectively.

^(b) Time to reach 50% activity was taken from Figs. 32 and 33.

F. Random Cassette Mutagenesis of Residues 197 through 228

Plasmid pΔ222 (Wells, et al. (1985) Gene 34, 315-323) was digested with PstI and BamHI and the 0.4 kb PstI/BamHI fragment (fragment 1, see Fig. 34) purified from a polyacrylamide gel by electroelution.

The 1.5 kb EcoRI/BamHI fragment from pS4.5 was cloned into M13mp9. Site directed mutagenesis was performed to create the A197 mutant and simultaneously insert a silent SstI site over codons 195-196. The mutant EcoRI/BamHI fragment was cloned back into pBS42. The pA197 plasmid was digested with BamHI and SstI and the 5.3 kb BamHI/SstI fragment (fragment 2) was purified from low melting agarose.

Complimentary oligonucleotides were synthesized to span the region from SstI (codons 195-196) to PstI (codons 228-230). These oligodeoxynucleotides were designed to (1) restore codon 197 to the wild type, (2) re-create a silent KpnI site present in pΔ222 at codons 219-220, (3) create a silent SmaI site over codons 210-211, and (4) eliminate the PstI site over codons 228-230 (see Fig. 35). Oligodeoxynucleotides were synthesized with 2% contaminating nucleotides at each cycle of synthesis, e.g., dATP reagent was spiked with 2% dCTP, 2% dGTP, and 2% dTTP. For 97-mers, this 2% poisoning should give the following percentages of non-mutant, single mutants and double or higher mutants per strand with two or more misincorporations per complimentary strand: 14% non-mutant, 28% single mutant, and 57% with ≥2 mutations, according to the general formula

$$f = \frac{\mu^n}{n!} e^{-\mu} .$$

where μ is the average number of mutations and n is a number class of mutations and f is the fraction of the total having that number of mutations. Complimentary oligodeoxynucleotide pools were phosphorylated and annealed (fragment 3) and then ligated at 2-fold molar excess over fragments 1 and 2 in a three-way ligation.

E. coli MM294 was transformed with the ligation reaction, the transformation pool-grown up over night and the pooled plasmid DNA was isolated. This pool represented 3.4×10^4 independent transformants. This plasmid pool was digested with PstI and then used to retransform E. coli. A second plasmid pool was prepared and used to transform B. subtilis (BG2036). Approximately 40% of the BG2036 transformants actively expressed subtilisin as judged by halo-clearing on casein plates. Several of the non-expressing transformants were sequenced and found to have insertions or deletions in the synthetic cassettes. Expressing BG2036 mutants were arrayed in microtiter dishes with 150μl of LB/12.5μg/mL chloramphenicol (cmp) per well, incubated at 37 °C for 3-4 hours and then stamped in duplicate onto nitrocellulose filters laid on LB 1.5% skim milk/5μg/mL cmp plates and incubated overnight at 33 °C (until halos were approximately 4-8 mm in diameter). Filters were then lifted to stacks of filter paper saturated with 1 x Tide commercial grade detergent, 50 mM Na₂CO₃, pH 11.5 and incubated at 65 °C for 90 min. Overnight growth plates were

Commassie stained and destained to establish basal levels of expression. After this treatment, filters were returned to pH7/skim milk/20μg/mL tetracycline plates and incubated at 37 °C for 4 hours to overnight.

Mutants identified by the high pH stability screen to be more alkaline stable were purified and analyzed for autolytic stability at high pH or high temperature. The double mutant C204/R213 was more stable than wild type at either high pH or high temperature (Table XXII).

This mutant was dissected into single mutant parents (C204 and R213) by cutting at the unique SmaI restriction site (Fig. 35) and either ligating wild type sequence 3' to the SmaI site to create the single C204 mutant or ligating wild type sequence 5' to the SmaI site to create the single R213 mutant. Of the two single parents, C204 was nearly as alkaline stable as the parent double mutant (C04/R213) and slightly more thermally stable. See Table XXII. The R213 mutant was only slightly more stable than wild type under both conditions (not shown).

Another mutant identified from the screen of the 197 to 228 random cassette mutagenesis was R204. This mutant was more stable than wild type at both high pH and high temperature but less stable than C204.

TABLE XXIIStability of subtilisin variants

Purified enzymes (200 μ g/mL) were incubated in 0.1M phosphate, pH 12 at 30°C for alkaline autolysis, or in 2mM CaCl₂, 50mM MOPS, pH 7.0 at 62°C for thermal autolysis. At various times samples were assayed for residual enzyme activity. Inactivations were roughly pseudo-first order, and t 1/2 gives the time it took to reach 50% of the starting activity in two separate experiments.

<u>Subtilisin variant</u>	t 1/2 (alkaline autolysis)		t 1/2 (thermal autolysis)	
	Exp. #1	Exp. #2	Exp. #1	Exp. #2
wild type	30	25	20	23
F50/V107/R213	49	41	18	23
R204	35	32	24	27
C204	43	46	38	40
C204/R213	50	52	32	36
L204/R213	32	30	20	21

G. Random Mutagenesis at Codon 204

Based on the above results, codon 204 was targeted for random mutagenesis. Mutagenic DNA cassettes (for codon at 204) all contained a fixed R213 mutation which was found to slightly augment the stability of the C204 mutant.

Plasmid DNA encoding the subtilisin mutant C204/R213 was digested with SstI and EcoRI and a 1.0 kb EcoRI/SstI fragment was isolated by electro-elution from polyacrylamide gel (fragment 1, see Fig. 35).

C204/R213 was also digested with SmaI and EcoRI and the large 4.7 kb fragment, including vector sequences and the 3' portion of coding region, was isolated from low melting agarose (fragment 2, see Fig. 36).

Fragments 1 and 2 were combined in four separate three-way ligations with heterophosphorylated fragments 3 (see Figs. 36 and 37). This heterophosphorylation of synthetic duplexes should preferentially drive the phosphorylated strand into the plasmid ligation product. Four plasmid pools, corresponding to the four ligations, were restricted with SmaI in order to linearize any single cut C204/R213 present from fragment 2 isolation, thus reducing the background of C204/R213. E. coli was then re-transformed with

Small-restricted plasmid pools to yield a second set of plasmid pools which are essentially free of C204/R213 and any non-segregated heteroduplex material.

These second enriched plasmid pools were then used to transform *B. subtilis* (BG2036) and the resulting four mutant pools were screened for clones expressing subtilisin resistant to high pH/temperature inactivation. Mutants found positive by such a screen were further characterized and identified by sequencing.

The mutant L204/R213 was found to be slightly more stable than the wild type subtilisin. See Table XXII.

Having described the preferred embodiments of the present invention, it will appear to those ordinarily skilled in the art that various modifications may be made to the disclosed embodiments, and that such modifications are intended to be within the scope of the present invention.

Claims

1. A subtilisin mutant derived by the substitution of at least one amino acid residue of a precursor subtilisin with a different amino acid, so that the subtilisin mutant has at least one property which is different from the same property of the precursor subtilisin, characterised by the substitution at one or more of Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214, and Gly215 of *Bacillus amyloliquefaciens* subtilisin and equivalent amino acid residues in other precursor subtilisins.
2. A subtilisin mutant having an amino acid sequence derived from the amino acid sequence of a precursor subtilisin by the substitution of more than one amino acid residue of said amino acid sequence of said precursor subtilisin by a different amino acid, so that the subtilisin mutant has at least one property which is different from the same property of the precursor subtilisin, characterized by substitutions at more than one of Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214, and Gly215 of *Bacillus amyloliquefaciens* subtilisin and equivalent amino acid residues in other precursor subtilisins, with the proviso that when substitution is made at any residue in the group Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 and Met222 a substitution is also made at at least one specified position not of that group.
3. The mutant of claim 2 wherein said combinations are selected from Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 and Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.
4. A subtilisin mutant derived by the deletion of one or more amino acid residues in a precursor subtilisin equivalent to 161-164 in *B. amyloliquefaciens* subtilisin, said deletion being made alone or in combination with substitutions in the amino acid sequence of the precursor subtilisin, and producing at least one property which is different from the same property of the precursor subtilisin.
5. A subtilisin mutant having altered substrate specificity when compared to a precursor subtilisin, the mutant being derived by the substitution of a different amino acid at the residue equivalent to Leu + 126 of *B. amyloliquefaciens* subtilisin, alone or in combination with other substitutions or deletions in the amino acid sequence of the precursor subtilisin.
6. A subtilisin mutant having altered substrate specificity when compared to a precursor subtilisin, the mutant being derived by the substitution of a different amino acid at the residue equivalent to Asp + 99 in *B. amyloliquefaciens* subtilisin, alone or in combination with other substitutions or deletions in the amino acid sequence of the precursor subtilisin.
7. A DNA sequence encoding the mutant of any one of the preceding claims.

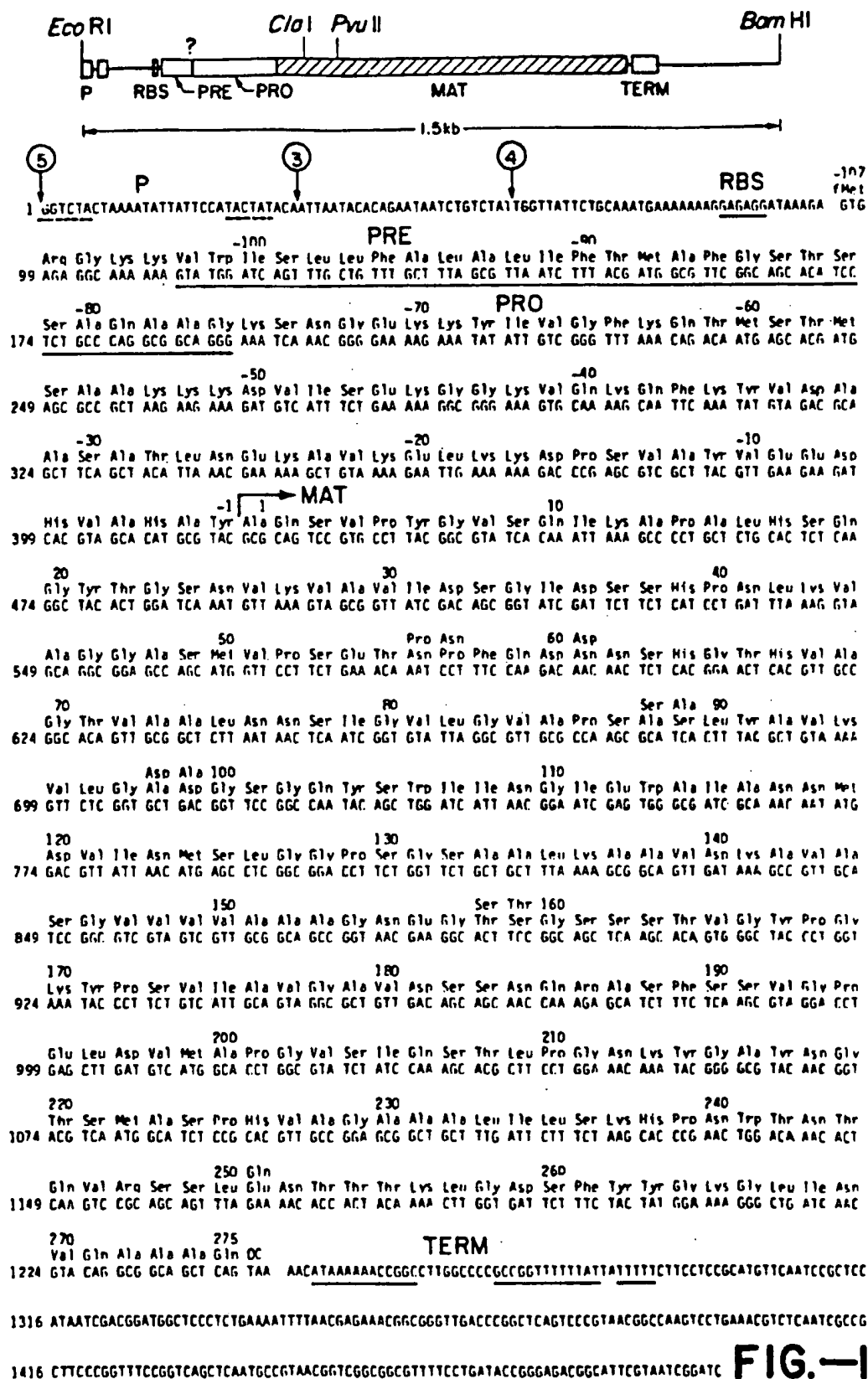
8. An expression vector containing the mutant DNA sequence of claim 7.
9. A host cell transformed with the expression vector or claim 8.

5 Patentansprüche

1. Subtilisinmutante, die durch Substitution zumindest eines Aminosäurerests eines Vorläufer-Subtilisins durch eine davon verschiedene Aminosäure hergeleitet ist, sodaß die Subtilisinmutante zumindest eine Eigenschaft aufweist, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet, gekennzeichnet durch die Substitution an einem oder mehreren von Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 und Gly215 von *Bacillus amyloliquefaciens*-Subtilisin und äquivalenten Aminosäureresten in anderen Vorläufer-Subtilisinen.
2. Subtilisinmutante mit einer Aminosäuresequenz, die aus der Aminosäuresequenz eines Vorläufer-Subtilisins durch Substitution mehr als eines Aminosäurerests der Aminosäuresequenz des Vorläufer-Subtilisins durch eine davon verschiedene Aminosäure hergeleitet ist, sodaß die Subtilisinmutante zumindest eine Eigenschaft aufweist, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet, gekennzeichnet durch Substitutionen an mehr als einem von Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 und Gly215 von *Bacillus amyloliquefaciens*-Subtilisin und äquivalenten Aminosäureresten in anderen Vorläufer-Subtilisinen, mit der Maßgabe, daß bei einer Substitution an irgendeinem Rest in der Gruppe Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 und Met222 eine Substitution auch an zumindest einer bestimmten Position durchgeführt wird, die nicht dieser Gruppe angehört.
3. Mutante nach Anspruch 2, worin die Kombinationen aus Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Tyr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 und Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217 ausgewählt sind.
4. Subtilisinmutante, die durch Löschung eines oder mehrerer Aminosäurereste in einem Vorläufer-Subtilisin, das 161-164 in *B. amyloliquefaciens*-Subtilisin äquivalent ist, hergeleitet ist, wobei die Löschung entweder alleine oder in Kombination mit Substitutionen in der Aminosäuresequenz des Vorläufer-Subtilisins erfolgt, und zumindest eine Eigenschaft ergibt, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet.
5. Subtilisinmutante mit geänderter Substratspezifität im Vergleich zu einem Vorläufer-Subtilisin, wobei die Mutante durch Substitution einer unterschiedlichen Aminosäure am Rest, der Leu + 126 von *B. amyloliquefaciens*-Subtilisin äquivalent ist, alleine oder in Kombination mit anderen Substitutionen oder Löschungen in der Aminosäuresequenz des Vorläufer-Subtilisins hergeleitet ist.
6. Subtilisinmutante mit geänderter Substratspezifität im Vergleich zu einem Vorläufer-Subtilisin, wobei die Mutante durch Substitution einer unterschiedlichen Aminosäure am Rest, der Asp + 99 im *B. amyloliquefaciens*-Subtilisin äquivalent ist, alleine oder in Kombination mit anderen Substitutionen oder Löschungen in der Aminosäuresequenz des Vorläufer-Subtilisins hergeleitet ist.
7. DNA-Sequenz, die für die Mutante nach einem der vorhergehenden Ansprüche kodiert.
8. Expressionsvektor, der die Mutanten-DNA-Sequenz von Anspruch 7 enthält.
9. Wirtszelle, die mit dem Expressionsvektor von Anspruch 8 transformiert ist.

Revendications

1. Mutant de subtilisine dérivé par la substitution d'au moins un résidu d'acide aminé d'une subtilisine précurseur et par un acide aminé différent de manière que le mutant de subtilisine ait au moins une propriété qui est différente de la même propriété de la subtilisine précurseur, caractérisé par la substitution à un ou plusieurs de Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 et Gly215 de la subtilisine de *Bacillus amyloliquefaciens* et les résidus d'acides aminés équivalents dans d'autres subtilisines précurseurs.
2. Mutant de subtilisine ayant une séquence d'acides aminés dérivée de la séquence d'acides aminés d'une subtilisine précurseur par la substitution de plus d'un résidu d'acide aminé de ladite séquence d'acides aminés de ladite subtilisine précurseur par un acide aminé différent de manière que le mutant de subtilisine ait au moins une propriété qui est différente de la même propriété de la subtilisine précurseur, caractérisé par des substitutions à plus d'un de Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 et Gly215 de la subtilisine de *Bacillus amyloliquefaciens* et des résidus d'acides aminés équivalents dans d'autres subtilisines précurseurs, à condition que quand la substitution est effectuée à tout résidu dans le groupe formé de Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 et Met222, une substitution soit également effectuée en au moins une position spécifiée ne faisant pas partie de ce groupe.
3. Mutant de la revendication 2 où lesdites associations sont choisies parmi Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 et Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.
4. Mutant de subtilisine dérivé par la délétion d'un ou plusieurs résidus d'acides aminés dans une subtilisine précurseur équivalente à 161-164 dans la subtilisine de *B. amyloliquefaciens*, ladite délétion étant effectuée seule ou en association avec des substitutions dans la séquence d'acides aminés de la subtilisine précurseur et la production d'au moins une propriété qui est différente de la même propriété de la subtilisine précurseur.
5. Mutant de subtilisine ayant une spécificité modifiée du substrat en comparaison avec une subtilisine précurseur, le mutant étant dérivé par la substitution d'un acide aminé différent au résidu équivalent à Leu + 126 de la subtilisine de *B. amyloliquefaciens*, seule ou en association avec d'autres substitutions ou délétions dans la séquence d'acides aminés de la subtilisine précurseur.
6. Mutant de subtilisine ayant une spécificité modifiée de substrat en comparaison avec une subtilisine précurseur, le mutant étant dérivé par la substitution d'un acide aminé différent au résidu équivalent à Asp + 99 dans la subtilisine de *B. amyloliquefaciens*, seule ou en association avec d'autres substitutions ou délétions dans la séquence d'acides aminés de la subtilisine précurseur.
7. Séquence d'ADN codant le mutant selon l'une quelconque des revendications précédentes.
8. Vecteur d'expression contenant la séquence d'ADN du mutant de la revendication 7.
9. Cellule hôte transformée par le vecteur d'expression de la revendication 8.



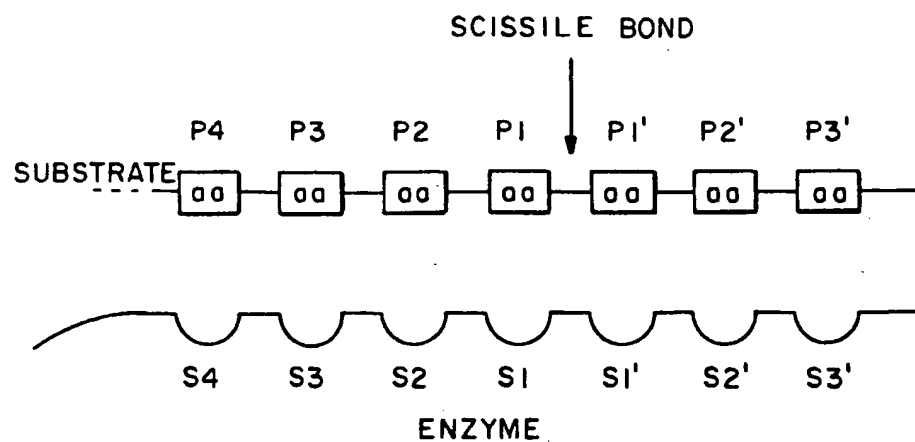


FIG.-2

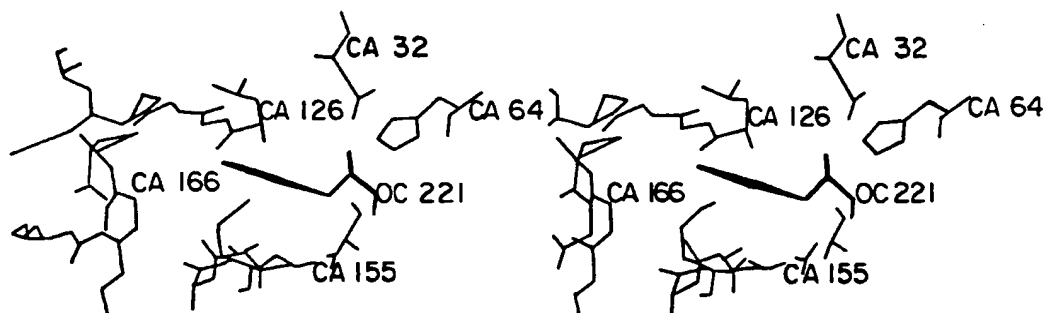


FIG.-3

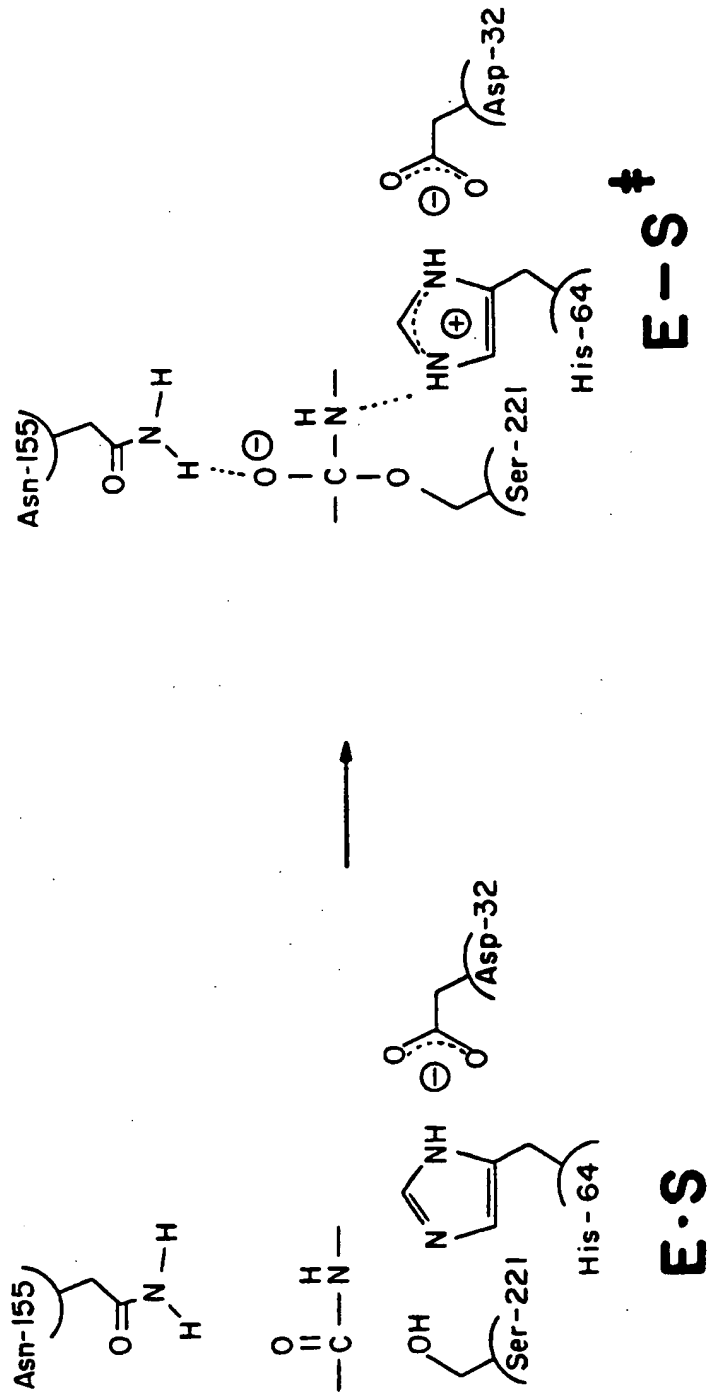


FIG.-4

Homology of *Bacillus proteases*

1. *Bacillus amyloliquifaciens*
 2. *Bacillus subtilis* var. 1169
 3. *Bacillus licheniformis* (carlsbergensis)

1	A	Q	S	V	P	Y	6	V	S	10	Q	I	K	A	P	A	L	H	S	Q	20
A	Q	S	V	P	Y	6	6	I	S	Q	I	I	K	A	P	A	L	H	S	Q	6
A	Q	T	V	P	Y	6		I	P	L	I	K	A	D	K	V	Q	A	Q	6	
21										30											40
Y	T	6	S	N	V	K		V	A	V	I	D	S	6	I	D	S	S	H	P	
Y	T	6	S	N	V	K		V	A	V	I	D	S	6	I	D	S	S	H	P	
F	K	6	A	N	V	K		V	A	V	L	D	T	6	I	Q	A	S	H	P	
41										50											60
D	L	K	V	A	6	6	A	S	M	V	P	S	E	T	N	P	F	Q	D		
D	L	N	V	R	6	6	A	S	F	V	P	S	E	T	N	P	Y	Q	D		
D	L	N	V	V	6	6	A	S	F	V	A	6	E	A	Y	N	T	.	D		
61										70											80
N	N	S	H	6	T	H	V	A	6	T	V	A	A	L	N	N	S	I	6		
6	S	S	H	6	T	H	V	A	6	T	I	A	A	L	N	N	S	I	6		
6	N	6	H	6	T	H	V	A	6	T	V	A	A	L	D	N	T	T	6		
81										90											100
V	L	6	V	A	P	S	A	S	L	Y	A	V	K	V	L	6	A	D	6		
V	L	6	V	S	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	6		
V	L	6	V	A	P	S	V	S	L	Y	A	V	K	V	L	N	S	S	6		
101										110											120
S	6	Q	Y	S	W	I	I	N	6	I	E	W	A	I	A	N	N	M	D		
S	6	Q	Y	S	W	I	I	N	6	I	E	W	A	I	S	N	N	M	D		
S	6	S	Y	S	6	I	V	S	6	I	E	W	A	T	T	N	6	M	D		

FIG.—5A—1

121									130								140
V	I	N	M	S	L	G	G	P	S	G	S	A	A	L	K	A	A
V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	V
V	I	N	M	S	L	G	G	A	S	G	S	T	A	M	K	Q	A
141									150								160
K	A	V	A	S	G	V	V	V	V	A	A	A	G	N	E	G	T
K	A	V	S	S	G	I	V	V	A	A	A	A	G	N	E	G	S
N	A	Y	A	R	G	V	V	V	V	A	A	A	G	N	S	G	N
161									170								180
S	S	S	T	V	G	Y	P	G	K	Y	P	S	V	I	A	V	G
S	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G
S	T	N	T	I	G	Y	P	A	K	Y	D	S	V	I	A	V	G
181									190								200
D	S	S	N	Q	R	A	S	F	S	S	V	G	P	E	L	D	V
N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L	D	V
D	S	N	S	N	R	A	S	F	S	S	V	G	A	E	L	E	V
201									210								220
P	G	V	S	I	Q	S	T	L	P	G	N	K	Y	G	A	Y	N
P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N
P	G	A	G	V	Y	S	T	Y	P	T	N	T	Y	A	T	L	N
221									230								240
S	M	A	S	P	H	V	A	G	A	A	A	L	I	L	S	K	H
S	M	A	T	P	H	V	A	G	A	A	A	L	I	L	S	K	H
S	M	A	S	P	H	V	A	G	A	A	A	L	I	L	S	K	H
241									250								260
W	T	N	T	Q	V	R	S	S	L	E	N	T	T	T	K	L	G
W	T	N	A	Q	V	R	D	R	L	E	S	T	A	T	Y	L	G
L	S	A	S	Q	V	R	N	R	L	S	S	T	A	T	Y	L	G
261									270								
F	Y	Y	G	K	G	L	I	N	V	Q	A	A	A	Q			
F	Y	Y	G	K	G	L	I	N	V	Q	A	A	A	Q			
F	Y	Y	G	K	G	L	I	N	V	E	A	A	A	Q			

FIG.—5A—2

ALIGNMENT OF B.AMYLOLIQUIFACIENS SUBTILISIN AND THERMITASE
1.B.amyloliquifaciens subtilisin
2.thermitase

[illegible]

FIG.—5B—1

A	A	A	G	N	E	B	T	S	150	S	S	S	T	U	B	Y	P	B	170
A	A	A	S	N	A	G	N	T	A	P	N	Y	P	A	K
Y	P	S	U	I	A	U	G	A	180	U	D	B	S	N	O	R	A	S	190
Y	S	N	A	I	A	U	A	S	T	D	D	N	D	N	K	S	S	F	S
S	U	G	P	E	L	D	U	M	200	A	P	G	U	S	I	O	S	T	210
T	Y	G	S	U	U	D	U	A	A	P	B	S	M	I	Y	S	T	Y	P
G	N	K	Y	S	A	Y	N	G	220	T	S	M	A	S	P	H	U	A	230
T	S	T	Y	A	S	L	S	G	T	S	M	A	T	P	H	U	A	G	A
A	A	L	I	L	S	K	H	P	240	N	U	T	N	T	O	U	R	S	250
A	G	L	L	A	S	O	B	R	S	.	.	A	S	N	I	R	A	A	L
E	N	T	T	T	K	.	L	G	260	D	S	F	Y	Y	G	K	G	L	I
E	N	T	A	D	K	I	S	G	T	G	T	Y	U	A	K	B	R	U	N
270	U	Q	A	A	A	D													
A	Y	K	A	U	D	Y													

FIG.—5B-2

TOTALLY CONSERVED RESIDUES IN SUBTILISINS

1	P	20
21	.	.	G	D	.	G	H	.	40
41	G	V	60
61	.	.	.	H	G	T	H	80
81	.	.	G	U	L	.	.	.	100
101	S	G	120
121	L	G	140
141	G	G	N	160
161	Y	P	U	.	.	.	180
181	S	F	200
201	P	G	G	T	220
221	S	M	A	.	P	H	U	A	G	240
241	R	260
261	N	280

FIG.—5C

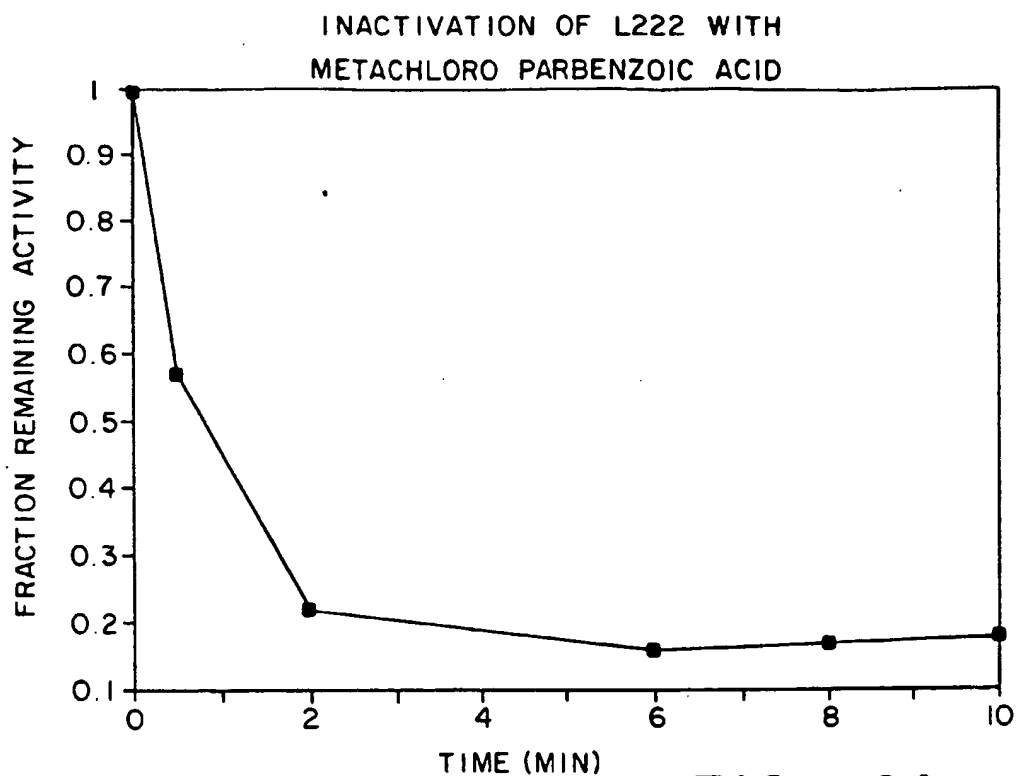


FIG.-6A

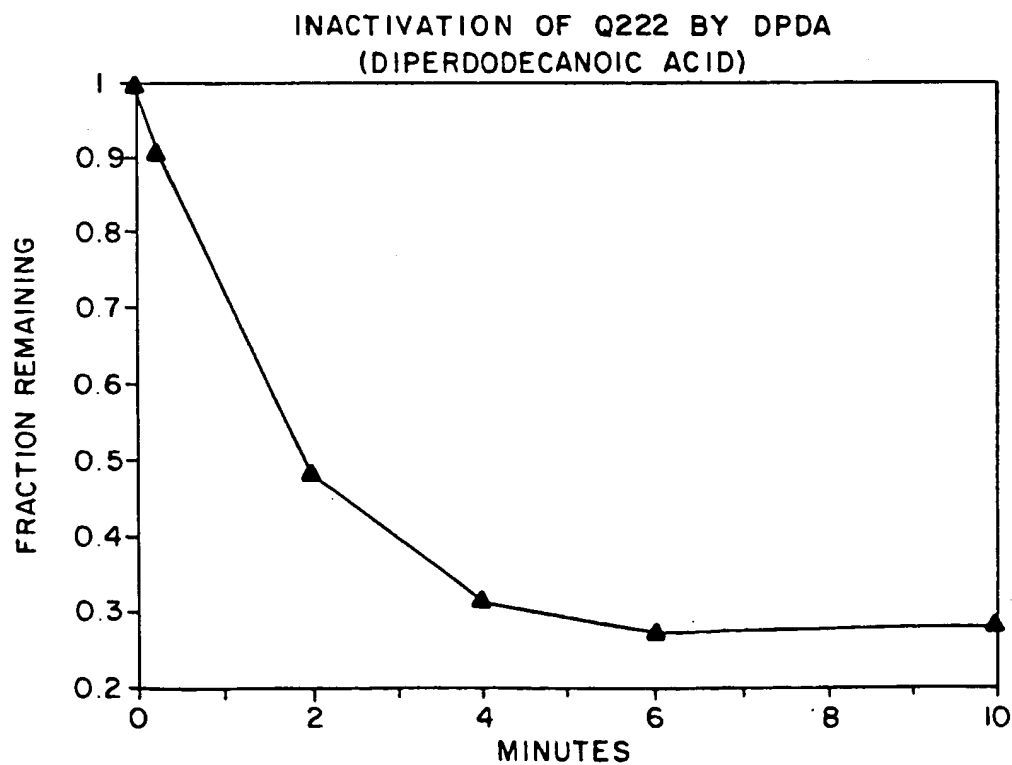


FIG.-6B

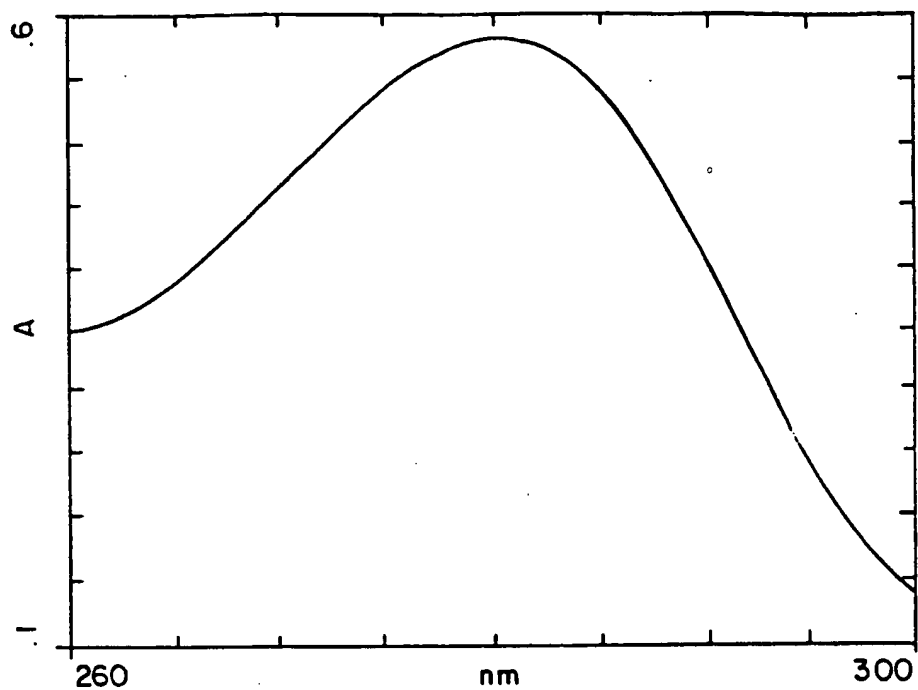


FIG.—7A

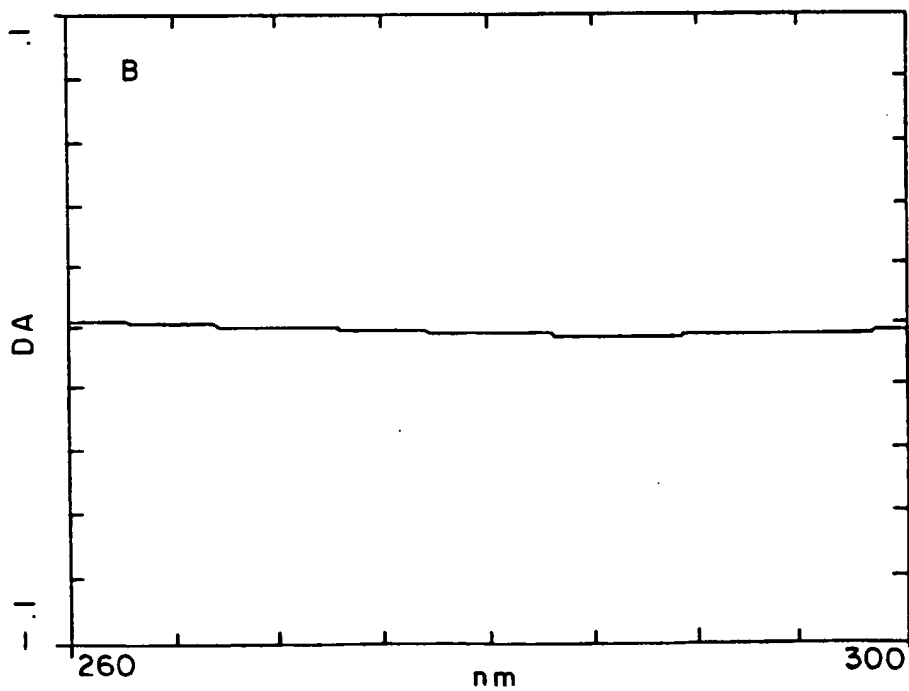


FIG.—7B

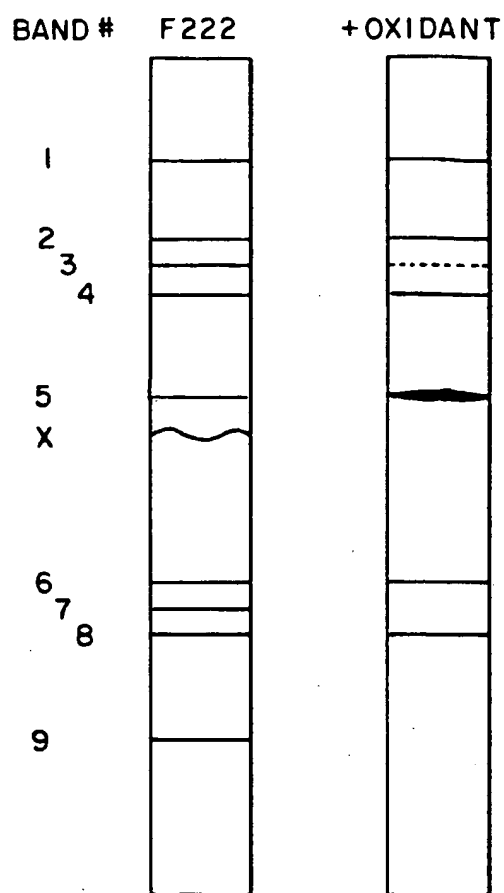


FIG.- 8

CNBr FRAGMENT MAP OF F222 MUTANT

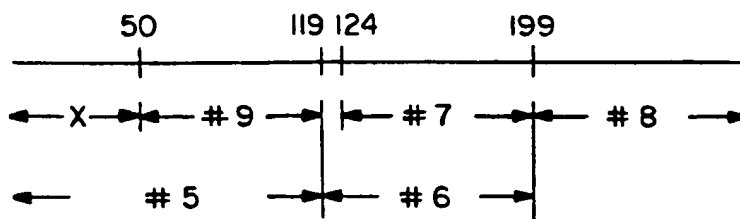


FIG.- 9

- FIG. 10**

1. Codon number: 117 120 124 126 130
2. Wild type amino acid sequence: Asn-Asn-Met-Asp-Val-Ile-Asn-Met-Ser-Leu-Gly-Gly-Pro-Ser
3. Wild type DNA sequence: 5' -AAC-AAT-ATG-GAC-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGA-CCT-TCT
TTG-TTA-TAC-CTG-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCT-GGA-AGA-5'
4. pΔ124:

5' -AAC-AAT-ATG-GAT-ATC-----C-GGG-GGC-CCT-TCT

TTG-TTA-TAC-CTA-TAG-----G-CCC-CCG-GGA-AGA-5'

Eco RV *Apa* I

*

5' -AAC-AAT-ATG-GAT-----C-GGG-GGC-CCT-TCT

TTG-TTA-TAC-CTA-TAG-----G-CCC-CCG-GGA-AGA-5'

Eco RV *Apa* I
5. pΔ124 cut with Eco RV and *Apa* I

*

5' -AAC-AAT-ATG-GAT-----C-GGG-GGC-CCT-TCT

TTG-TTA-TAC-CTA-TAG-----G-CCC-CCG-GGA-AGA-5'

Eco RV *Apa* I

*

5' -AAC-AAT-ATG-GAT-----C-GGG-GGC-CCT-TCT

TTG-TTA-TAC-CTA-TAG-----G-CCC-CCG-GGA-AGA-5'

Eco RV *Apa* I
6. Cut pΔ124 ligated with cassettes:

*

5' -AAC-AAT-ATG-GAT-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGC-CCT-TCT

TTG-TTA-TAC-CTA-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCG-GGA-AGA-5'

*

5' -AAC-AAT-ATG-GAT-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGC-CCT-TCT

TTG-TTA-TAC-CTA-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCG-GGA-AGA-5'
7. Mutagenesis primer for pΔ124::

5' -AAC-AAT-ATG-GAT-ATC-C-GGG-GGC-CCT-TCT-GGT-TC-3'

5' -AAC-AAT-ATG-GAT-ATC-C-GGG-GGC-CCT-TCT-GGT-TC-3'
8. Mutants made: 1124, L124 AND C126

FIG.—II

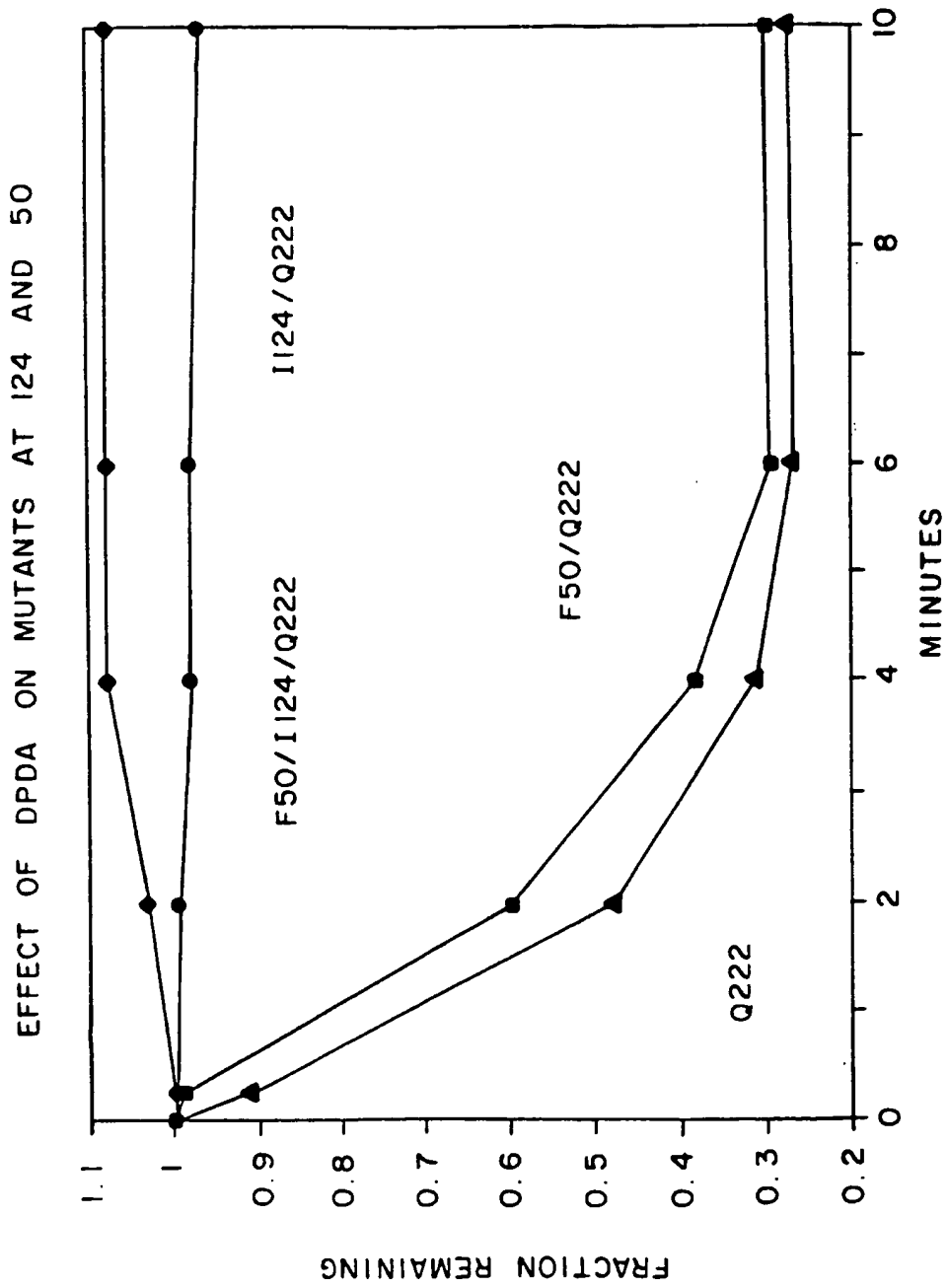


FIG.-12

166

Thr Ser Gly Ser Ser Thr Val Gly Tyr Pro Gly

1. Wild type DNA sequence:

5'-ACT TCC GGC AGC TCA AGC ACA GTG GGC TAC CCT GGT-3'
3'-TGA AGG CCG TCG AGT TCG TGT CAC CCG ATG GGA CCA-5'

2. pΔ166 DNA sequence:

$$\begin{array}{ccccccc}
 \star & & & & & & \star \\
 5' \text{--ACT} & \text{TCC} & \text{GGG} & \text{AGC} & \text{TCA} & \text{A} & \text{---C CCG GGT--3'} \\
 3' \text{--TGA} & \text{AGG} & \text{CCC} & \text{TGC} & \text{AGT} & \text{T} & \text{---G GGC CCA--5'} \\
 & & & \text{SacI} & & & \text{XmaI}
 \end{array}$$

3. pΔ166 cut with SacI and XmaI:

5'-ACT TCC GGG AGC T
3'-TGA AGG CCCp

4. Cut pΔ166 ligated with duplex DNA cassette pools:

5'-ACT TCC GGG AGC TCA AGC ACA GTG NNN TAC CCG GGT-3' ★
3'-TGA AGG CCC TCG AGT TCG TGT CAC NNN ATG GGC CCA-5' ★★ ★

MUTAGENESIS PRIMER 37 MER

5' AA GGC ACT TCC GGG AGC TCA ACC CGG GTA AA TAC CCT 3'

FIG. -13

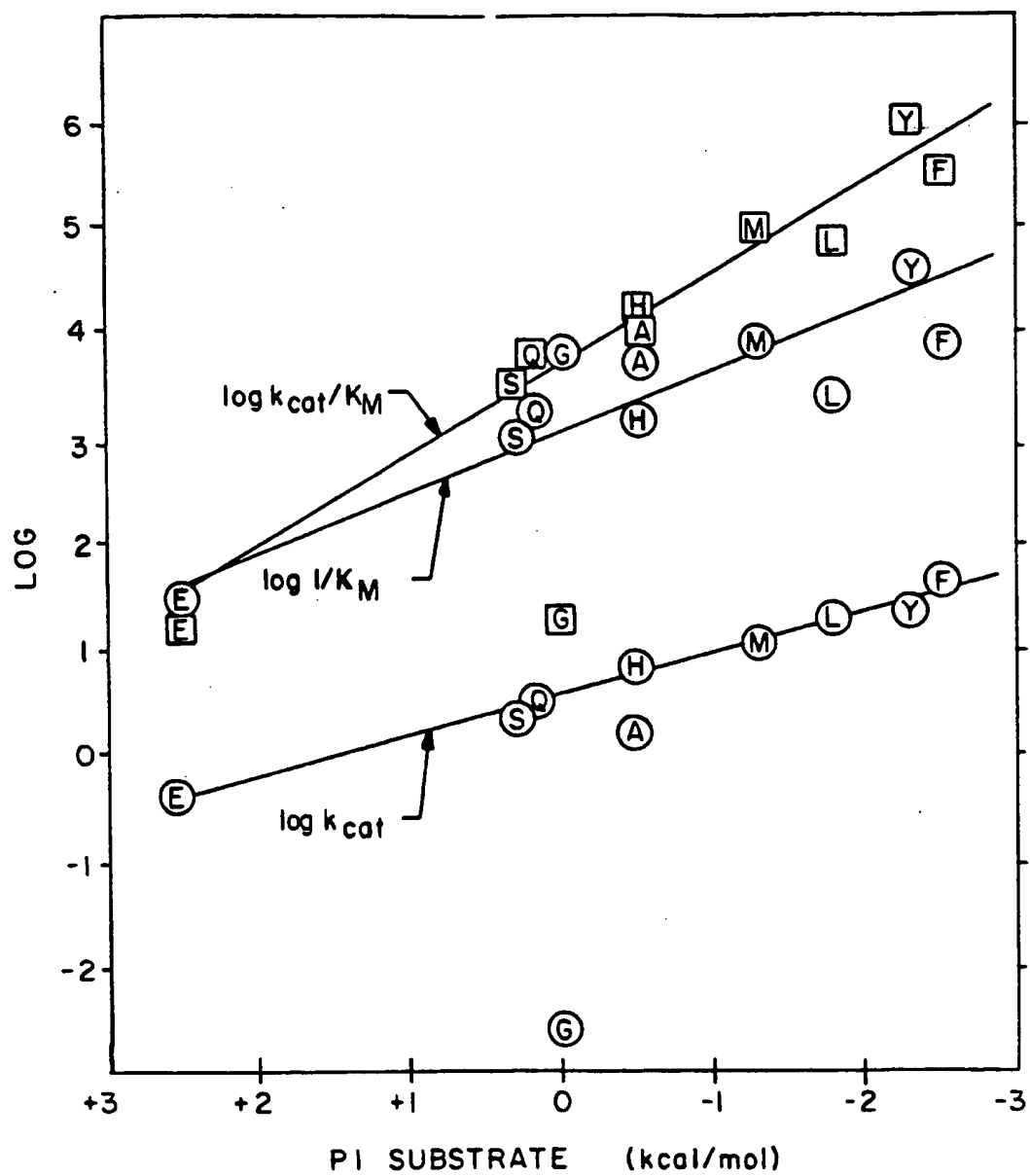


FIG. - 14

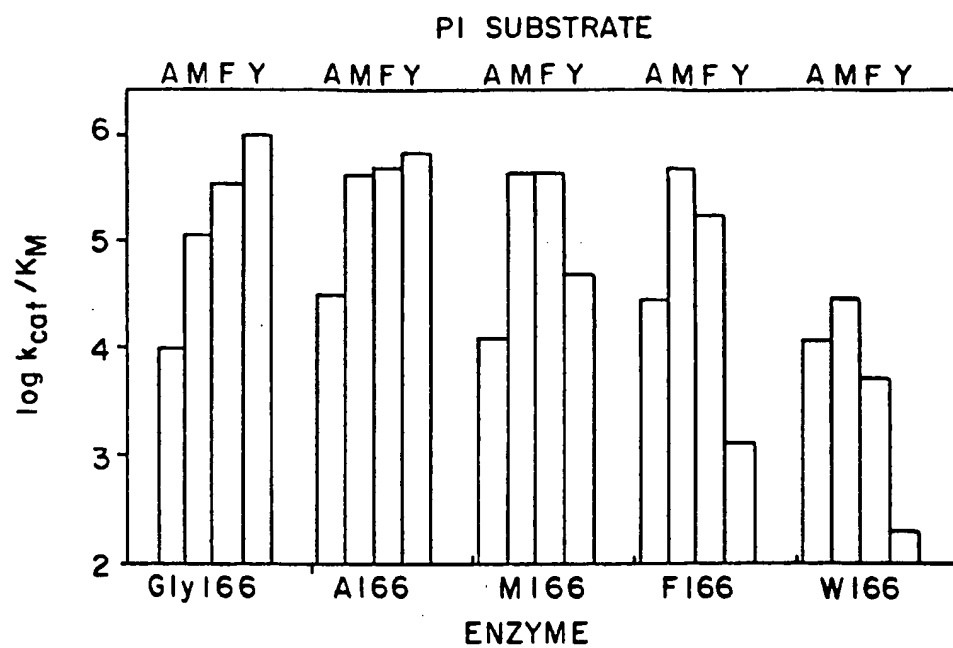


FIG. -15A

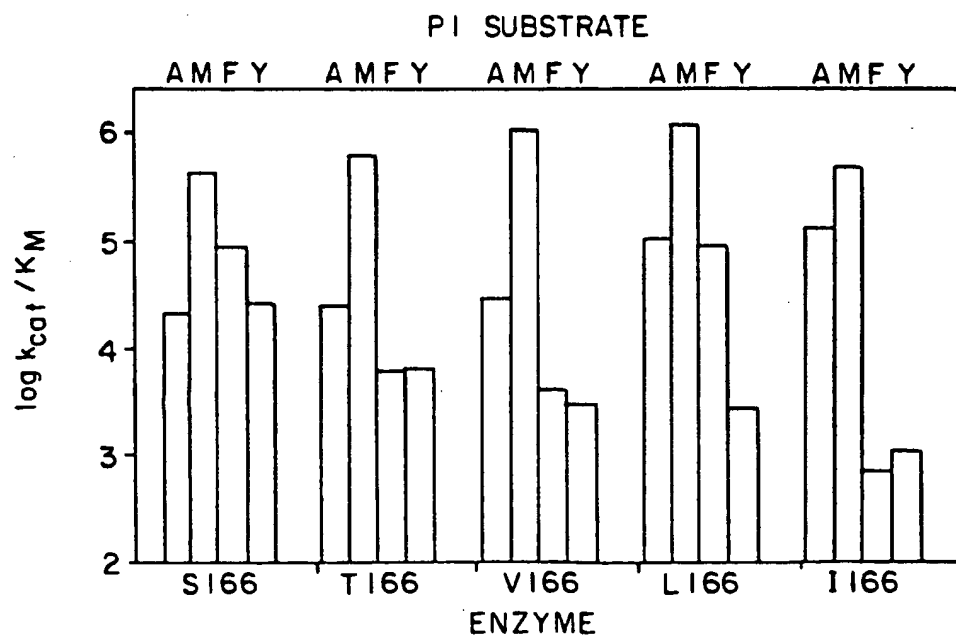


FIG. -15B

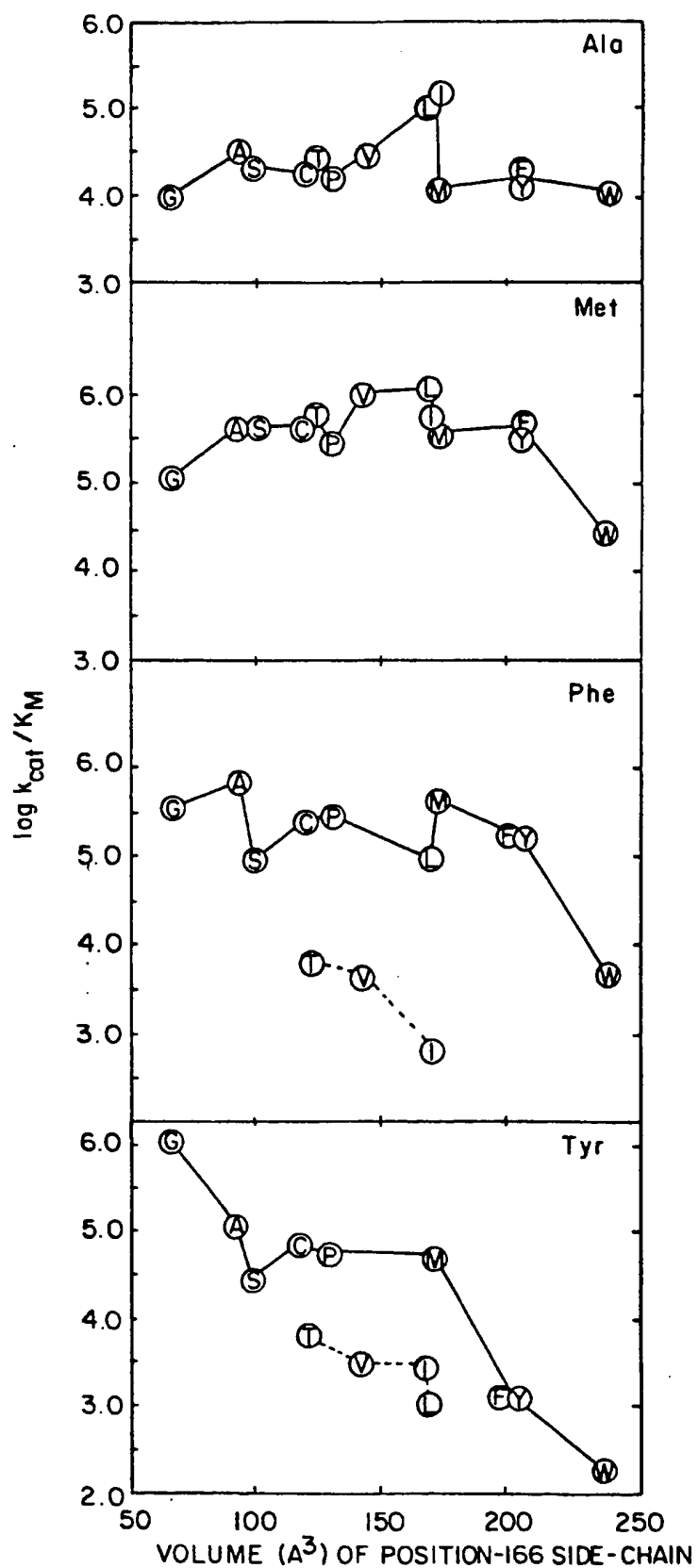


FIG.-16

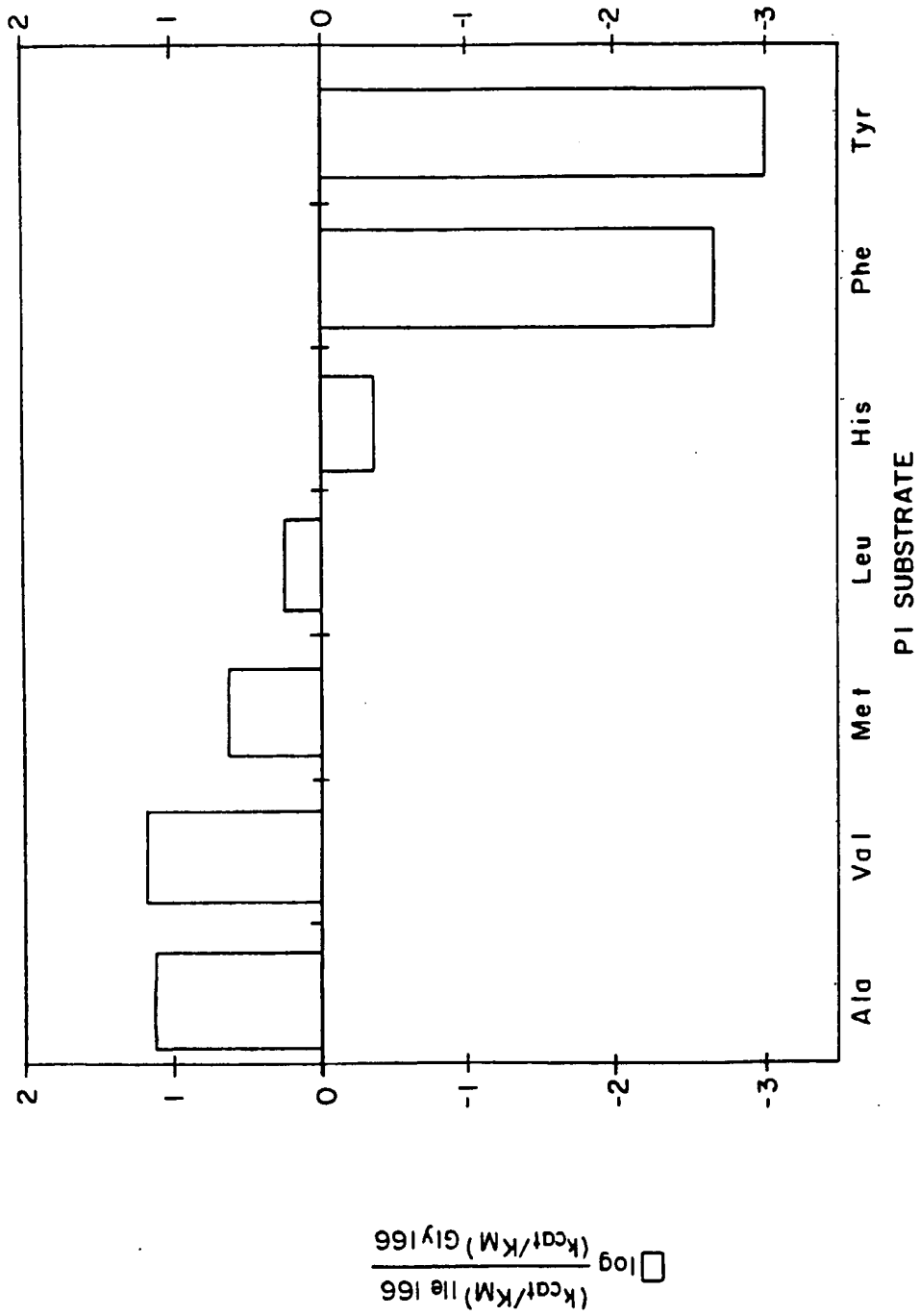


FIG. - 17

GLY-169 CASSETTE MUTAGENESIS

WILD TYPE AMINO ACID SEQUENCE:		CODON:		162		169		173			
				SER	SER	THR	VAL	GLY	TYR	PRO	SER
1.	WILD TYPE DNA SEQUENCE	5'	TCA AGC ACA GTG GGC TAC CCT GGT AAA TAC CCT TCT	3'	AGT TCG TGT CAC CCG ATG GGA CCA TTT ATG GGA AGA						
2.	P169 DNA SEQUENCE	5'	TCA AGC ACA GTC GGG TAC CCT-----GA TAT CCT TCT	3'	AGT TCG TGT CAC GCC ATG GGA CT ATA GGA AGA						
											ECORV
											KPN1
3.	P169 CUT WITH KPN1 AND ECORV:	5'	TAC AGC ACA GTC GGG TAC	3'	AGT TCG TGT CAC CCP						
											PAT CCT TCT
											TA GGA AGA
4.	CUT P169 LIGATED WITH OLIGONUCLEOTIDE POOLS	5'	TAC AGC ACA GTG GGG TAC CCT NNN AAA TAT CCT TGT	3'	AGT TCG TGT CAC CCC ATG GGA NNN TTT ATA GGA AGA						
	MUTAGENESIS PRIMER FOR P169	5'	AAG CAC AGT GGG GTA CCC TGA TAT CCT TCT GTC A	3'							

FIG.-18

1. Codon number: 100 104 105 108
2. Wild type amino acid sequence: Gly-Ser-Gly-Gln-Tyr-Ser-Trp-Ile-Ile-
3. Wild type DNA sequence: 5'-GGT-TCC-GGC-CAA-TAC-AGC-TGG-ATC-ATT-3'

C-AGC-TGG

Aul I
4. Primer for *Hind* III
 Insertion at 104:

GCTT

Hind III
5. Primers for 104 mutants:

GCTT

Hind III
6. Mutants made: A, M, L, S, AND H104

FIG.—19

1. Codon number: 148 150 152 155
2. Wild type amino acid sequence: Val-Val-Val-Ala-Ala-Ala-Gly-Asn-Glu
3. Wild type DNA sequence: 5'-GTA-GTC-GTT-GCG-GCA-GCC-GGT-AAC-GAA-3'

4. V152/P153 5'-GTA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3'

GTA-CCC

* *

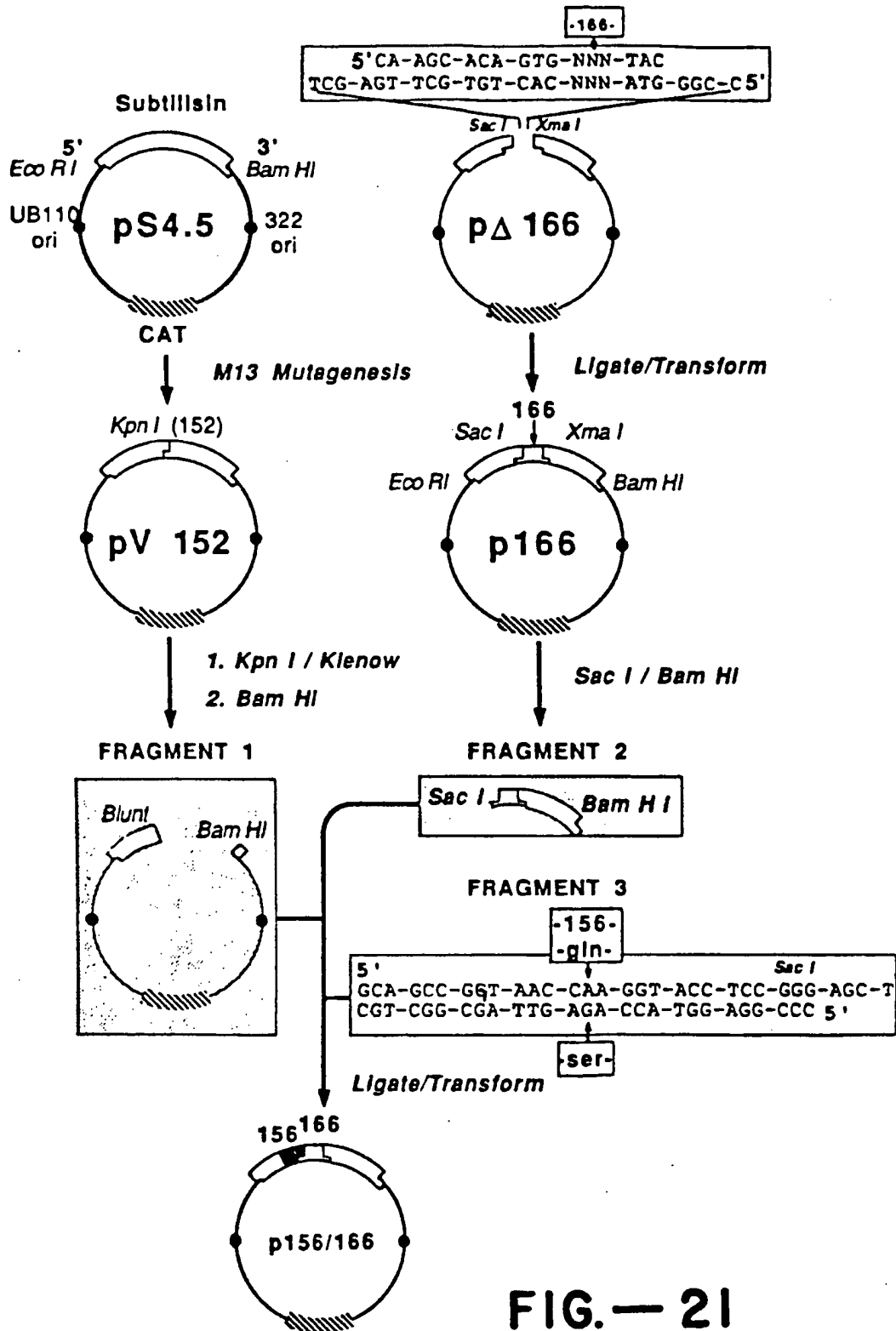
Kpn I

5. S 152: 5'-GTA-GTC-GTT-GCG-AGC-GCC-GGT-AAC-GAA-3'

6. G 152: 5'-GTA-GTC-GTT-GCG-GGC-GCC-GGT-AAC-GAA-3'

**

FIG.—20



1. Codon number: 211 215 217 220
2. Wild type amino acid sequence: Gly-Asn-Lys-Tyr-Gly-Ala-Tyr-Asn-Gly-Thr-Ser-Met-Ala
3. Wild type DNA sequence: 5' -GGA-AAC-AAA-TAC-GGG-GCG-TAC-AAC-GGT-ACG-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCC-CGC-ATG-TTG-CCA-TGC-AGT-TAC-CGT-5'
4. pΔ217
5' -GGA-AAC-AAA-TAC-GGC-GCC-TAC-----GG-ATA-TGA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-CGG-ATG-----CC-TAT-AGT-TAC-CGT-5'
Nar I Eco RV
5. pΔ217 cut with Nar I and Eco RI
5' -GGA-AAC-AAA-TAC-GG*
CCT-TTG-TTT-ATG-CCG-Gp
* PA-TCA-ATG-GCA
T-AGT-TAC-CGT-5'
6. Cut pΔ217 ligated with cassettes:
5' -GGA-AAC-AAA-TAC-GGC-GCG-NNN-AAC-GGT-ACA-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-CGC-NNN-TTG-CCA-TGT-AGT-TAC-CGT-5'
* ** * * *
7. Mutagenesis primer for pΔ217:
5' -GA-AAC-AAA-TAC-GGC-GCC-TAC-GGA-TAT-CAA-TGG-CAT-3'
* * * *
8. Mutants made: All 19 at 217

FIG.—22

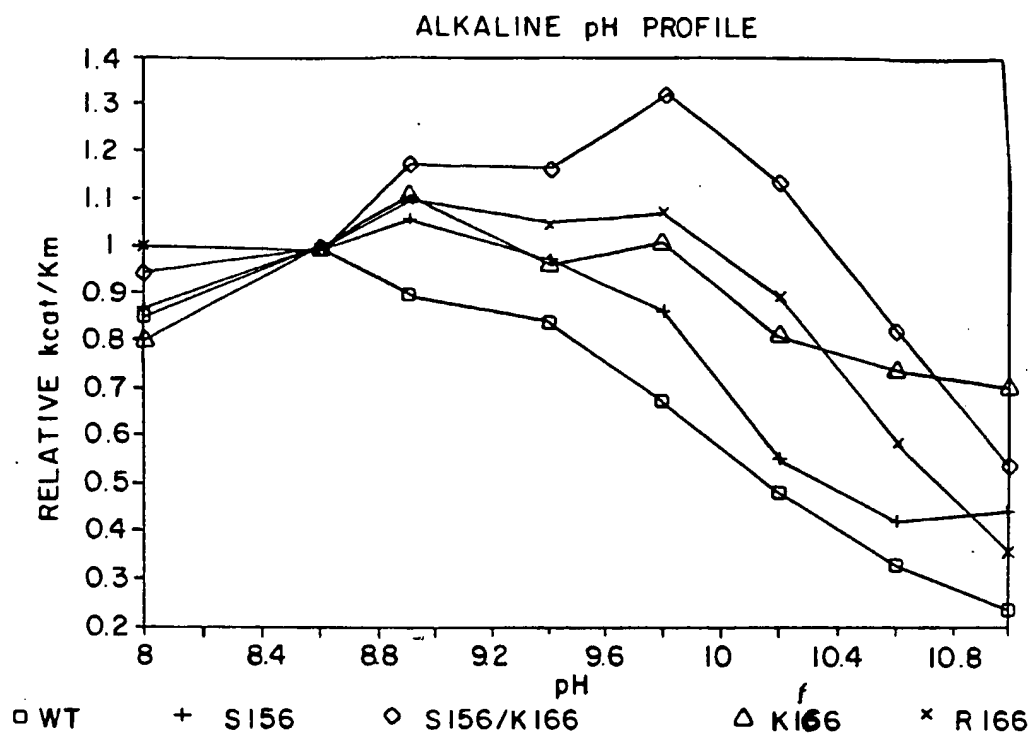


FIG. - 23A

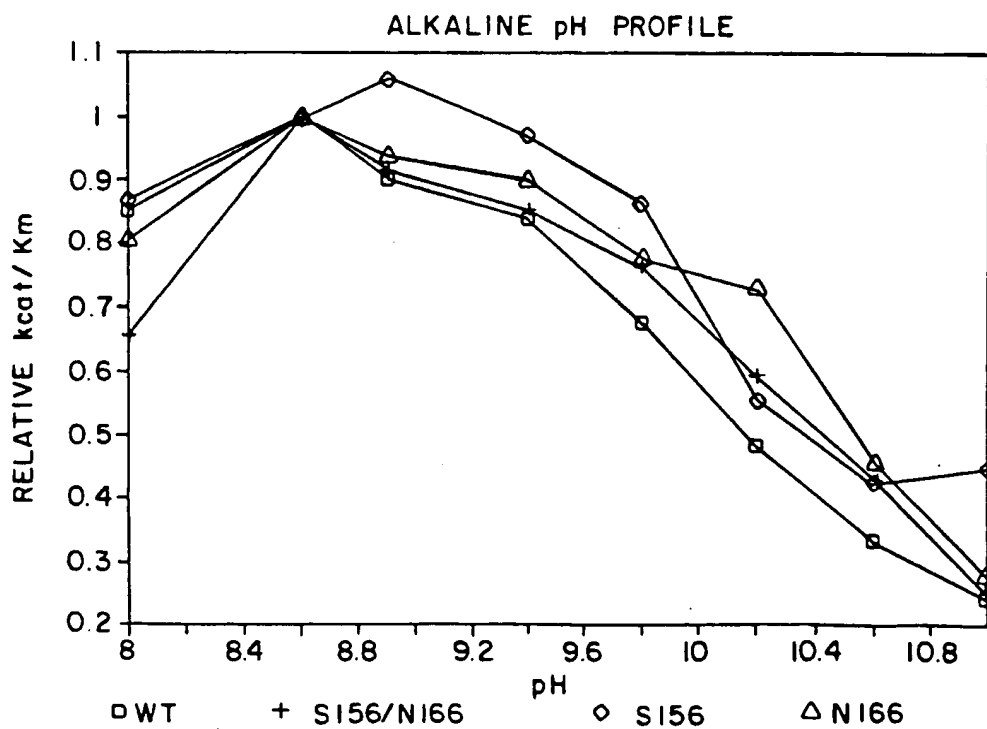


FIG. - 23B

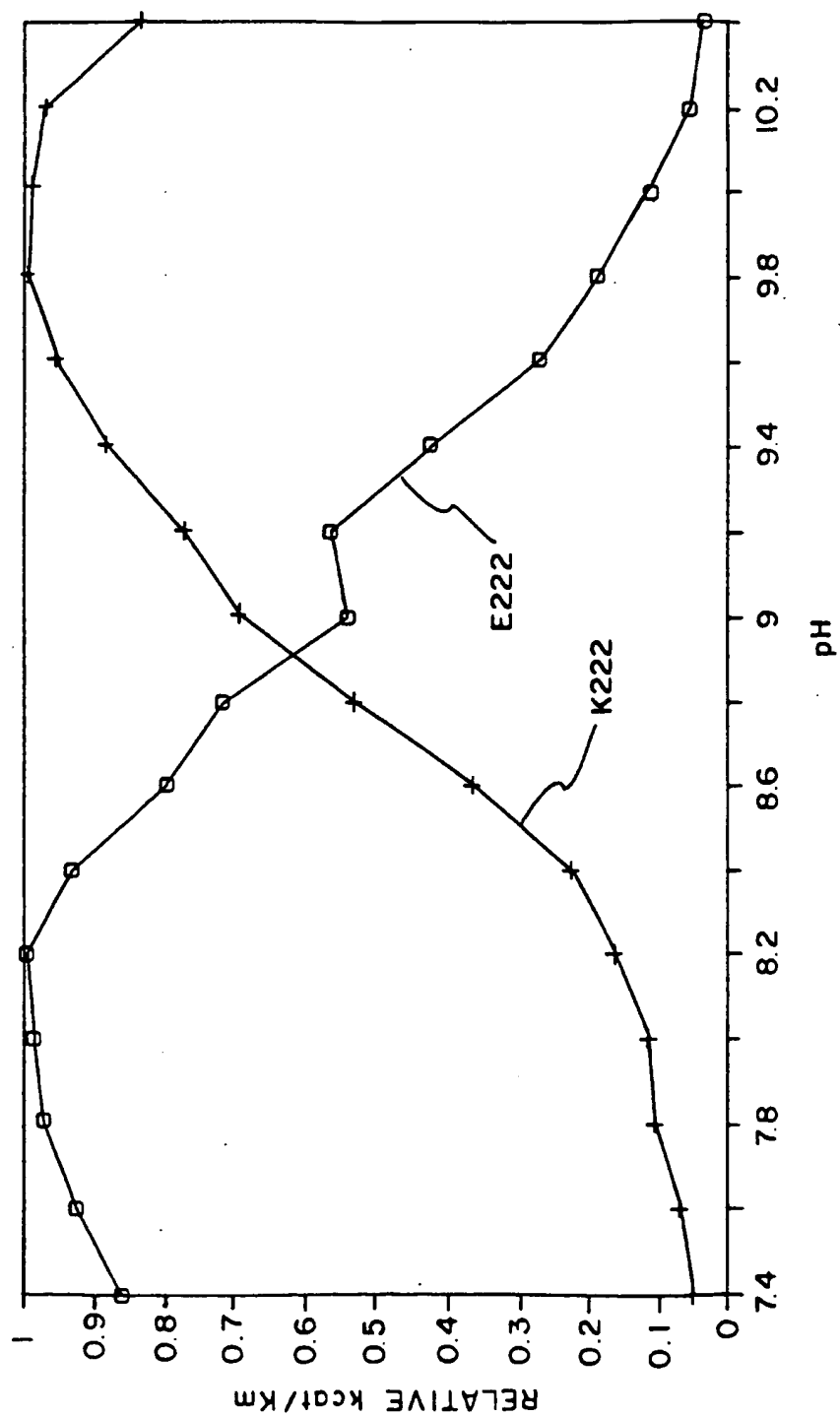


FIG. - 24

1. Codon number: 91 95 100
2. Wild type amino acid sequence: Tyr-Ala-Val-Lys-Val-Leu-Gly-Ala-Asp-Gly-Ser
3. Wild type DNA sequence: 5'-TAC-GCT-GTA-AAA-GTT-CTC-GGT-GCT-GAC-GGT-TCC
ATG-CGA-CAT-TTT-CAA-GAG-CCA-CTG-CCA-AGG-5'
4. pΔ95: 5'-TAC-GCG-T-^{★ ★}-----CTC-GCT-GCA-GAC-GGT-TCC
ATG-CGC-A-^{Mu I}-----GAG-CGA-CGT-CTG-CCA-AGG-5'
^{Pst I}
5. pΔ95 cut with *MuI* and *Pst I* 5'-TA [★] PGAC-GGT-TCC
ATG-CGCP A-CGT-CTG-CCA-AGG-5'
6. Cut pΔ95 ligated with cassettes: 5'-TAC-GCG-GTA-AAA-GTT-CTC-GGT-GCA-GAC-GGT-TCC
ATG-CGC-CAT-TTT-CAA-GAG-CCA-CGT-CTG-CCA-AGG-5'
[★]
7. Mutagenesis primer for pΔ95: 5'-CA-TCA-CTT-TAC-GCG-T-CTC-GCT-GCA-GAC-GGT-TCC
^{★ ★ ★ ★ ★}
8. Mutants made: C94, C95, D96

FIG.-25

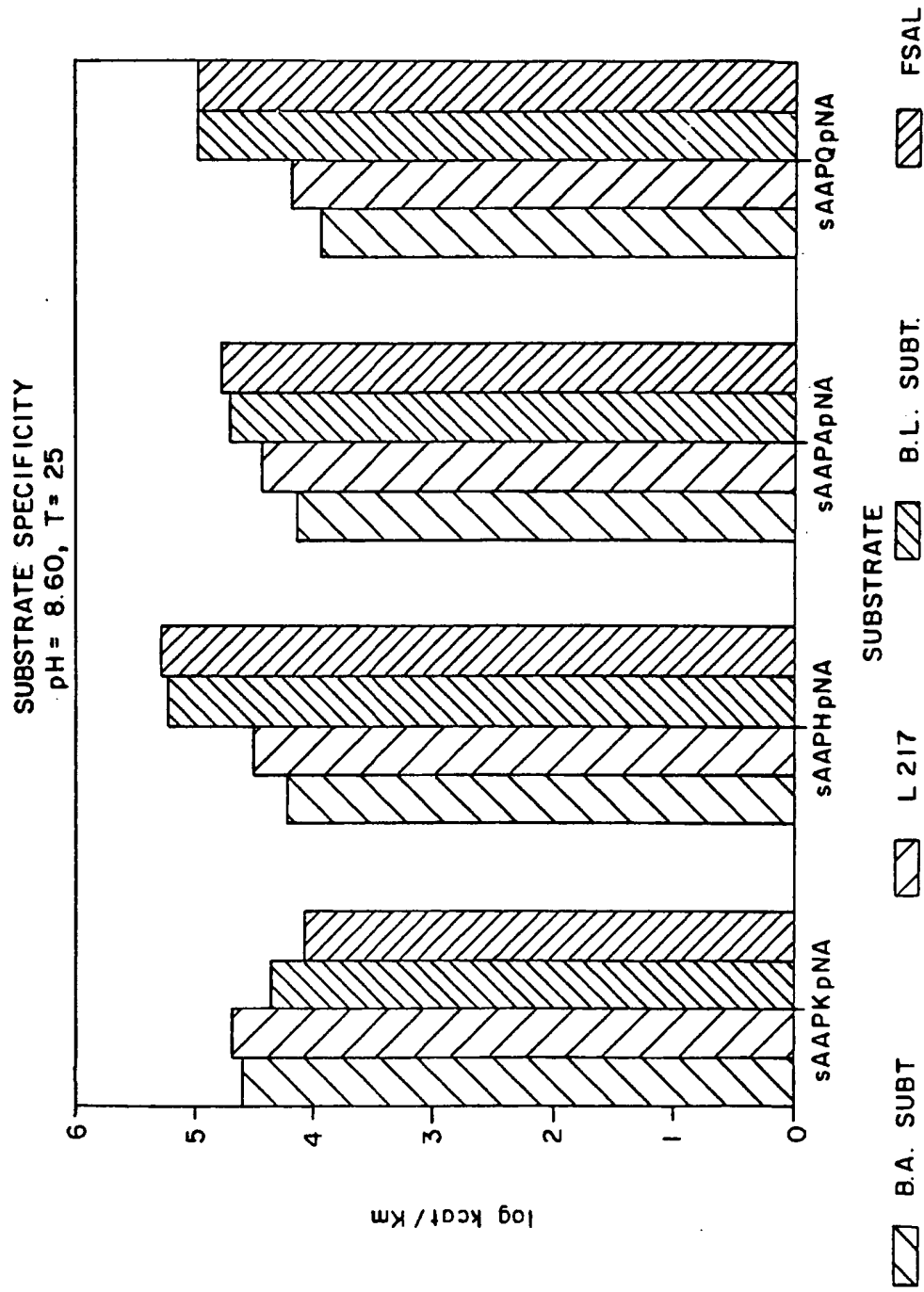


FIG.-26

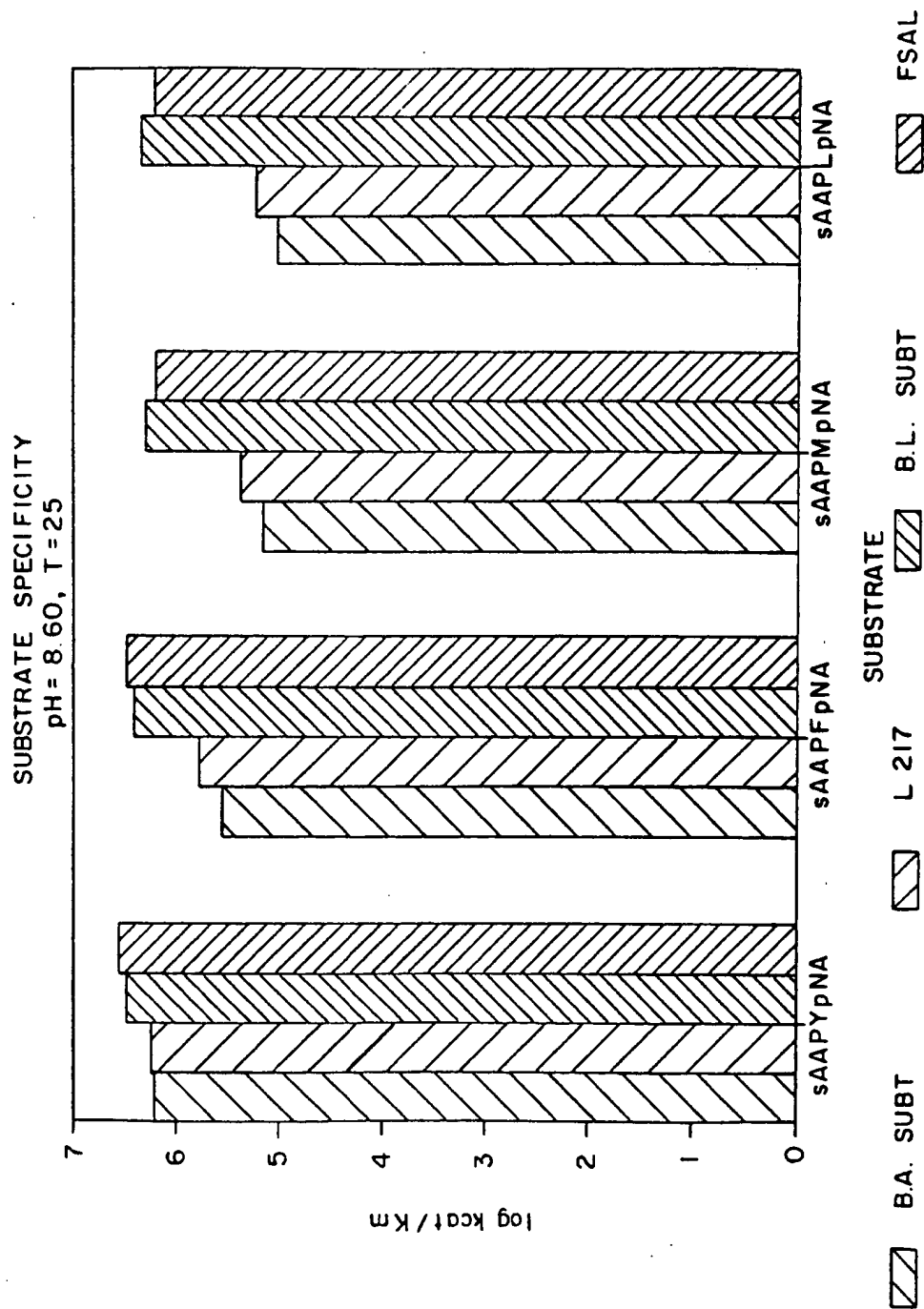


FIG.-27

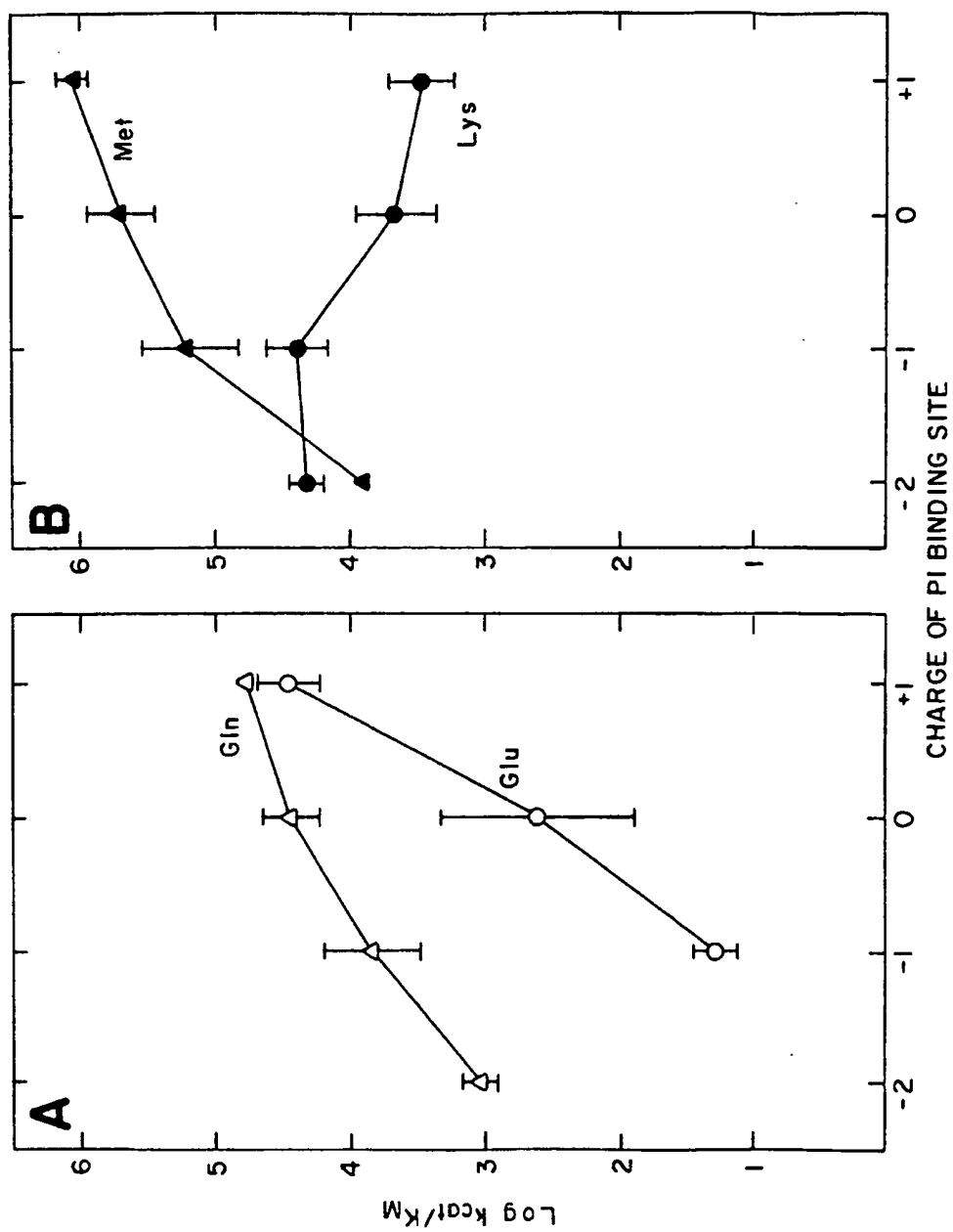


FIG.-28

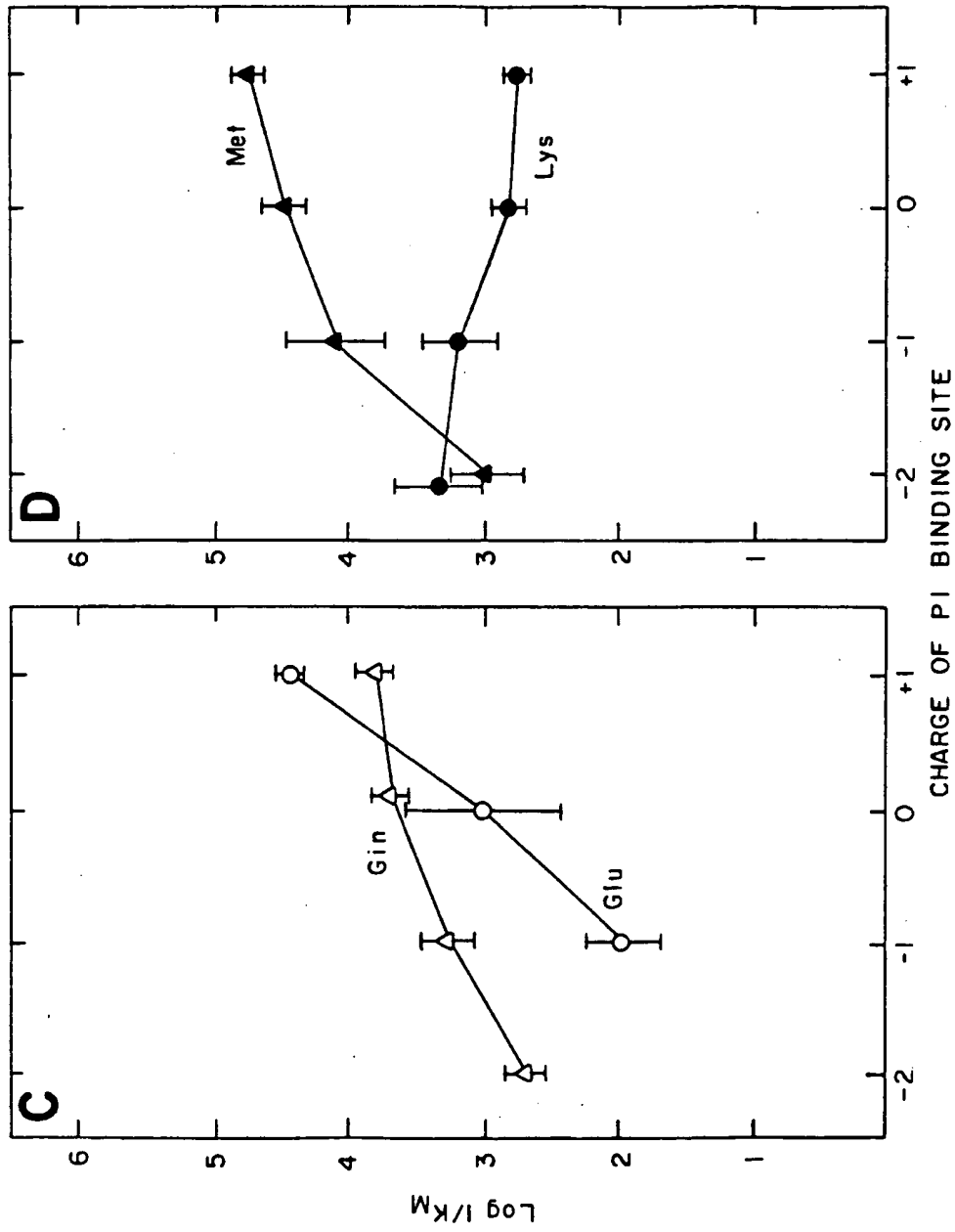


FIG.-28

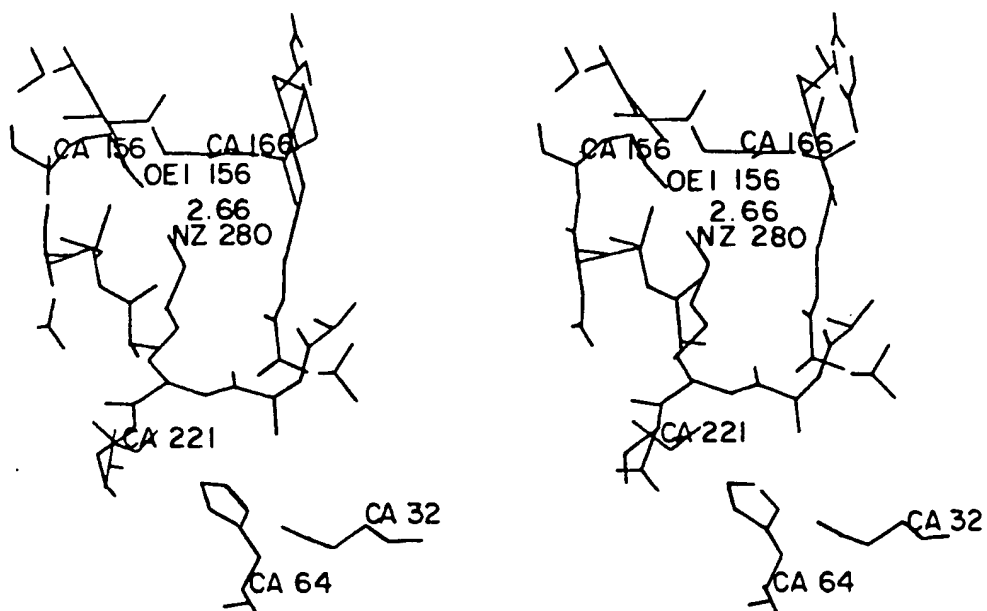


FIG. — 29A

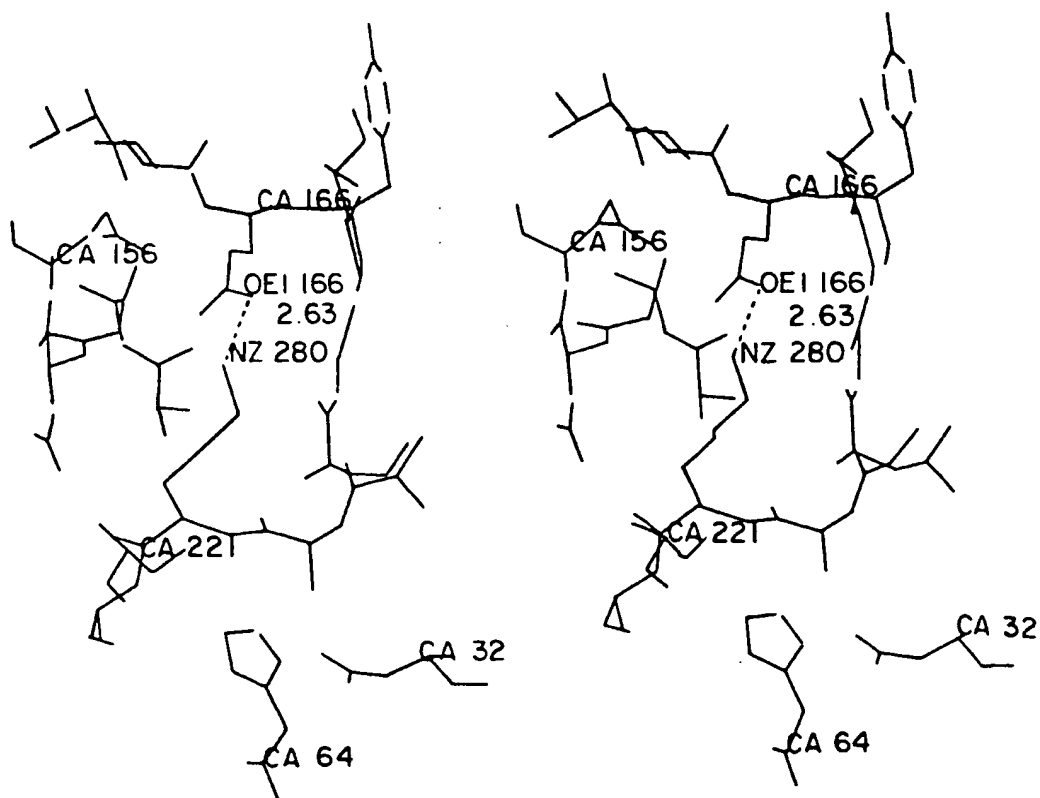


FIG. — 29B

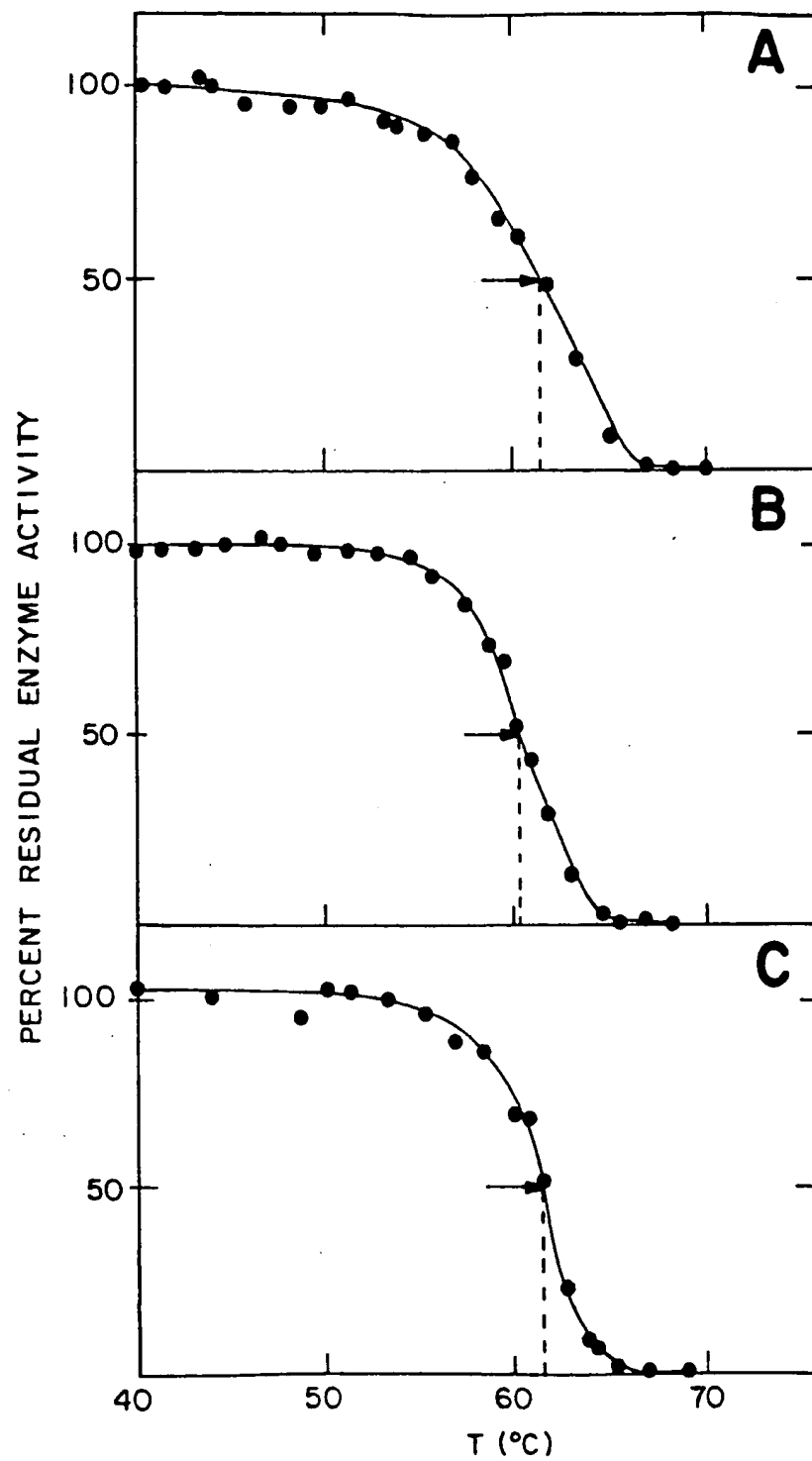


FIG.-30

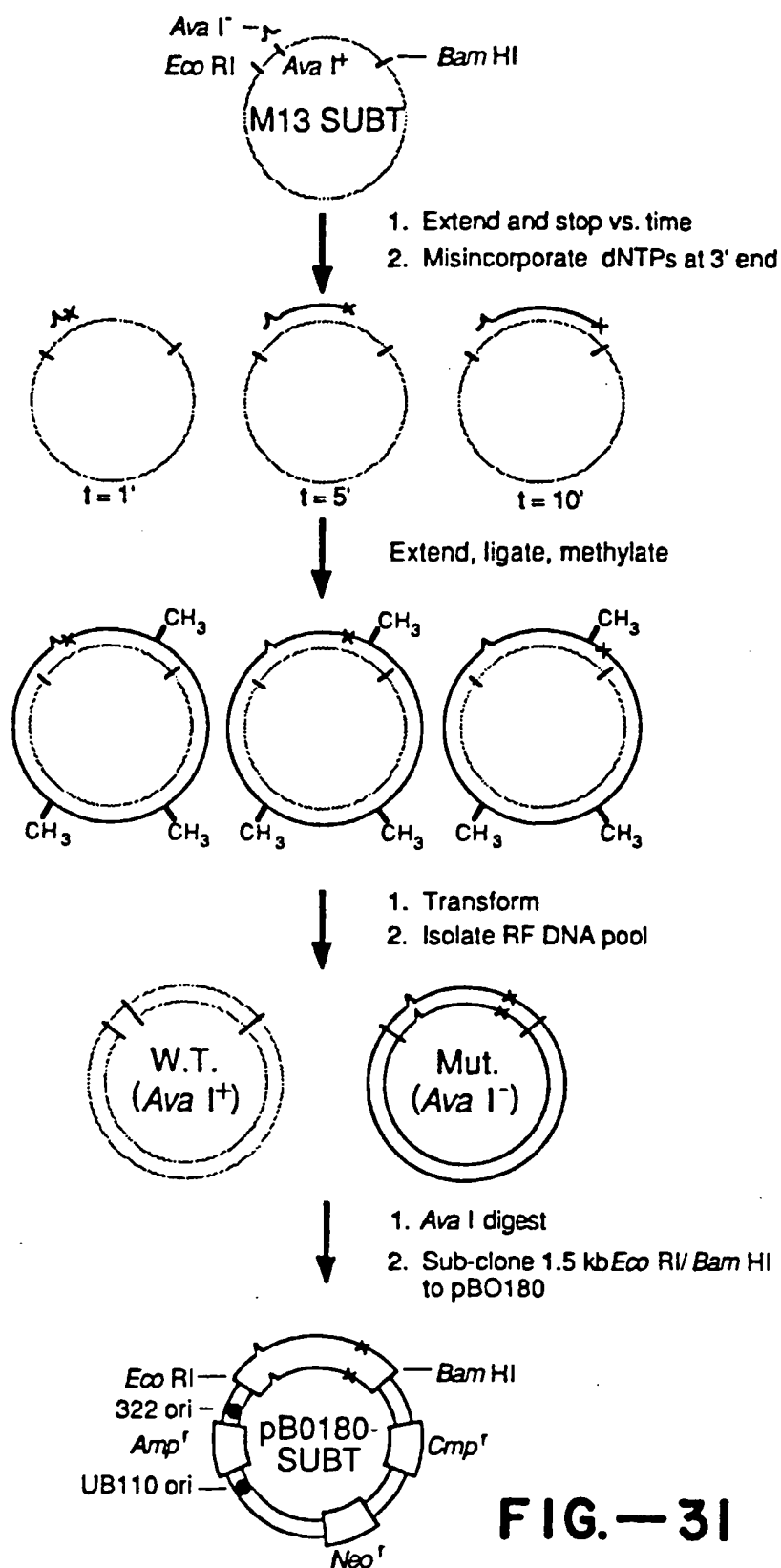


FIG.—31

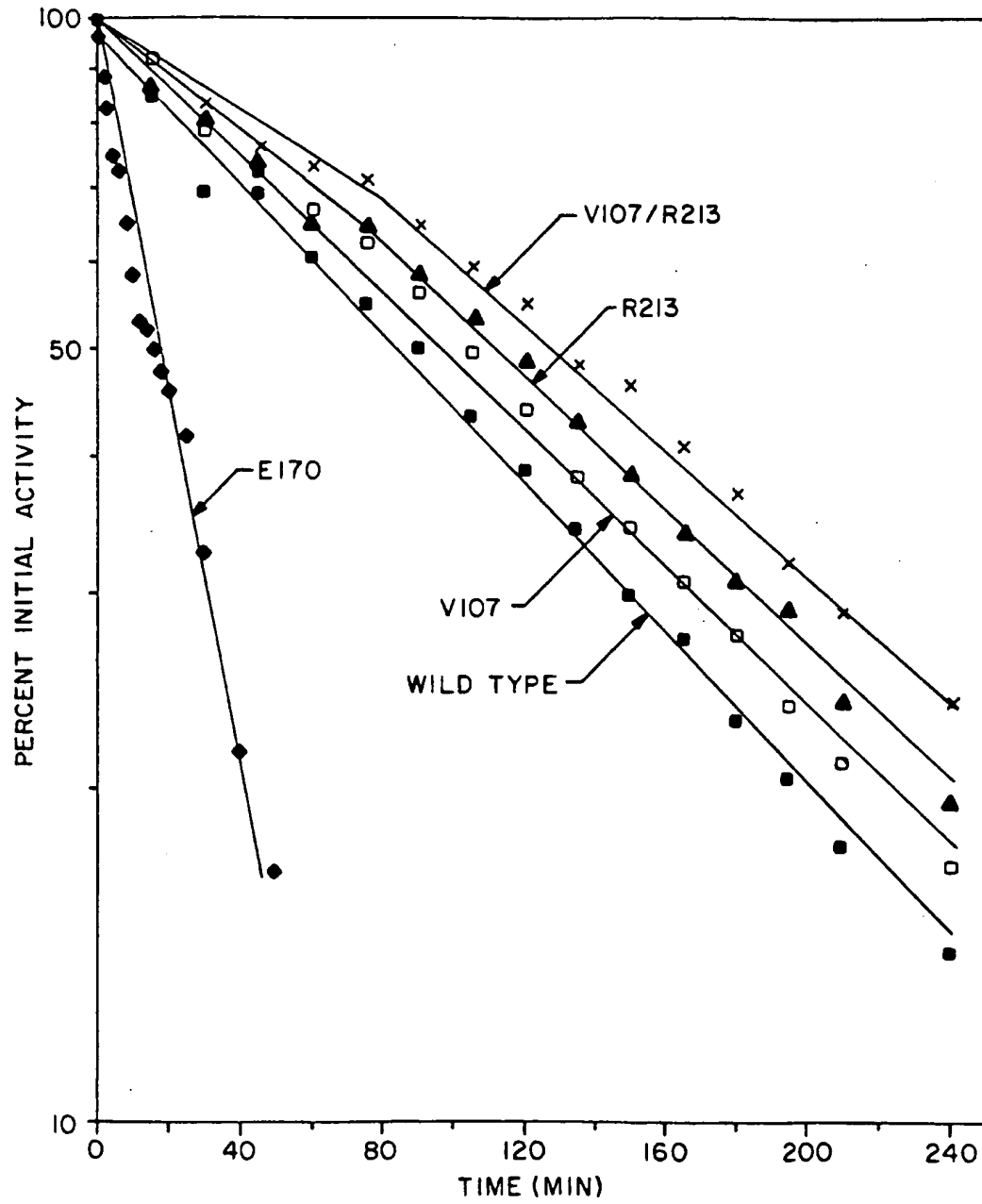


FIG. - 32

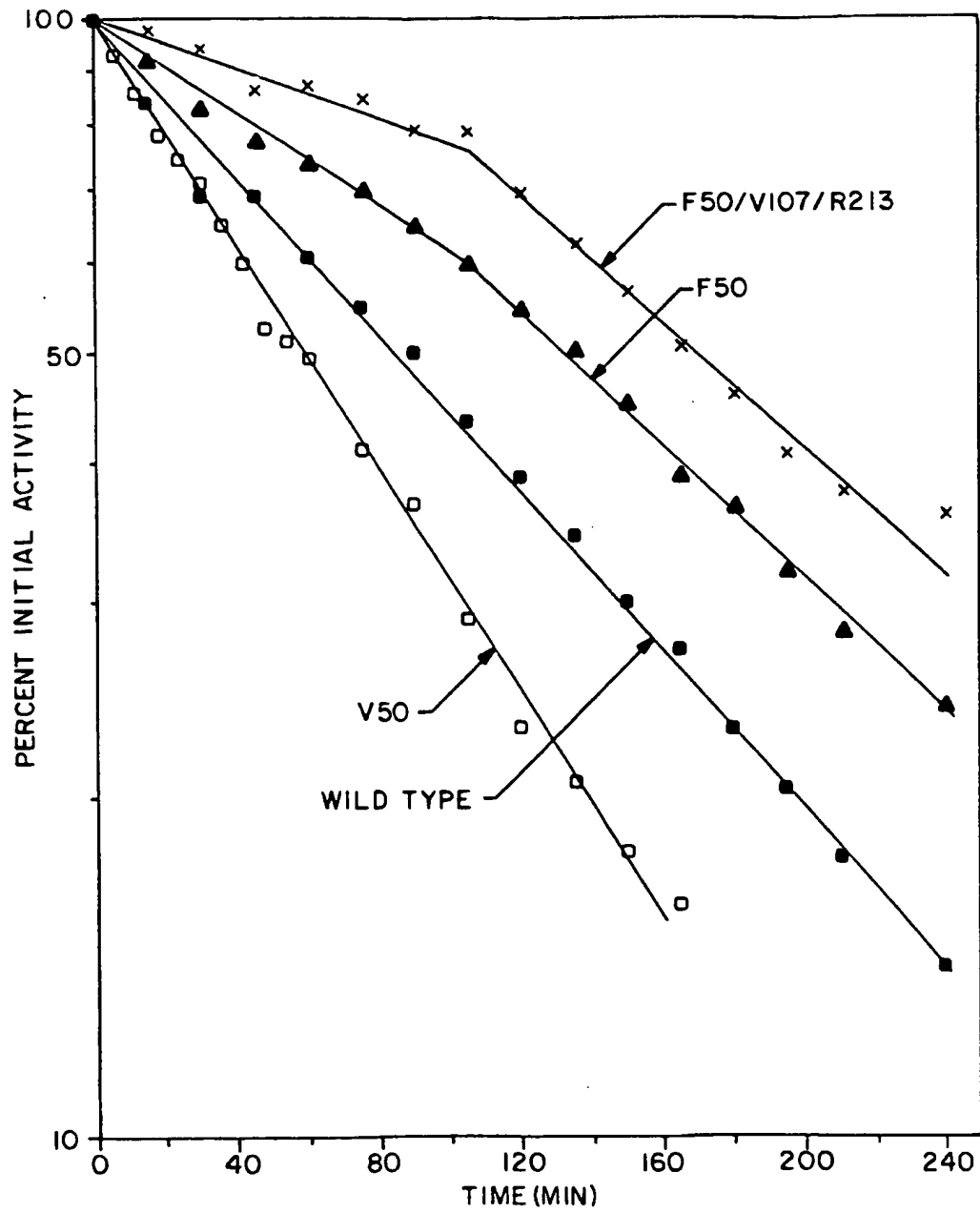


FIG.-33

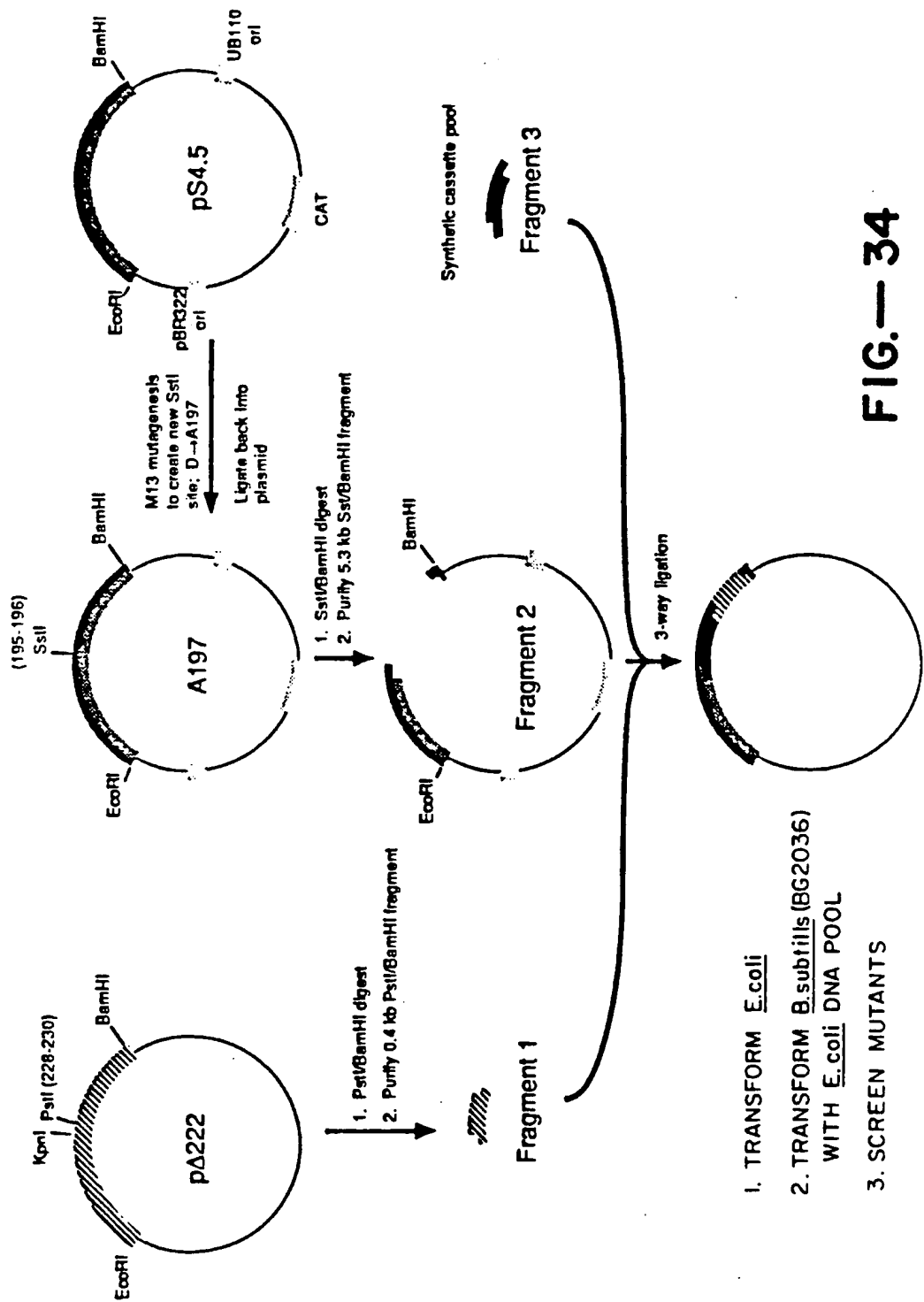


FIG.— 34

	195	200	206
W.T.A.A.:	Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Gln		
W.T. DNA:	GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
pΔ222 DNA:	GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
A197 DNA:	<u>GAG CTC</u> GCA GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAG CGT CAG TAC CGT GGA CCG CAT AGA TAG GTT		
	<i>Sst</i> I		
Fragments from pΔ222 and A197 cut w/ <i>Pst</i> I, <i>Sst</i> I:	GAG-CT		
	Cp		
	*		
pΔ222, A197 cut & ligated w/ oligodeoxy- nucleotide pools:	<u>GAG CTC GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA</u>		
	<u>CTC GAG CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT</u>		
	<i>Sst</i> I		
	*		
	207	210	218
W.T.A.A.:	Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn		
W.T. DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TGC TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
pΔ222 DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TGC TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
A197 DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TGC TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
	*	*	
Fragments from pΔ222 and A197 cut w/ <i>Pst</i> I, <i>Sst</i> I:	<u>AGC ACG CTT CCC GGG AAC AAA TAC GGG GCG TAC AAC</u>		
	<u>TGC TGC GAA GGG CCC TTG TTT ATG CCC CGC ATG TTG</u>		
	<i>Sma</i> I		
	219	220	230
W.T.A.A.:	Gly Thr Ser Met Ala Ser Pro His Val Ala Gly Ala		
W.T. DNA:	GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'		
	CCA TGC AGT TAC CGT AGA GGC GTG CAA CGG CCT CGC-5'		
pΔ222 DNA:	<u>GGT ACC</u> TCA-----CG CAC <u>GCT GCA</u> GGA GCG-3'		
	CCA TGG AGT-----GC GTG CGA CGT CCT CGC-5'		
	<i>Kpn</i> I	<i>Pst</i> I	
A197 DNA:	GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA GTG CCT CGC-5'		
		pGGA GCG-3'	
Fragments from pΔ222 and A197 cut w/ <i>Pst</i> I, <i>Sst</i> I:		A CGT CCT CGC-5'	
	*	*	
pΔ222, A197 cut & ligated w/ oligodeoxy- nucleotide pools:	<u>GGT ACC</u> TCA ATG GCA TCT CCG CAC GTT GCA GGA GCG-3'		
	<u>CCA TGG AGT TAC CGT AGA GGC GTG CAA</u> CGT CCT CGC-5'		
	<i>Kpn</i> I	<i>Pst</i> I destroyed	

Oligodeoxynucleotide pools synthesized with 2% contaminating nucleotides in each cycle to give
 -15% of pool with 0 mutations, -28% of pool with single mutations, and
 -57% of pool with 2 or more mutations, according to the general formula $f = \frac{\mu^n}{n!} e^{-\mu}$.

FIG.—35

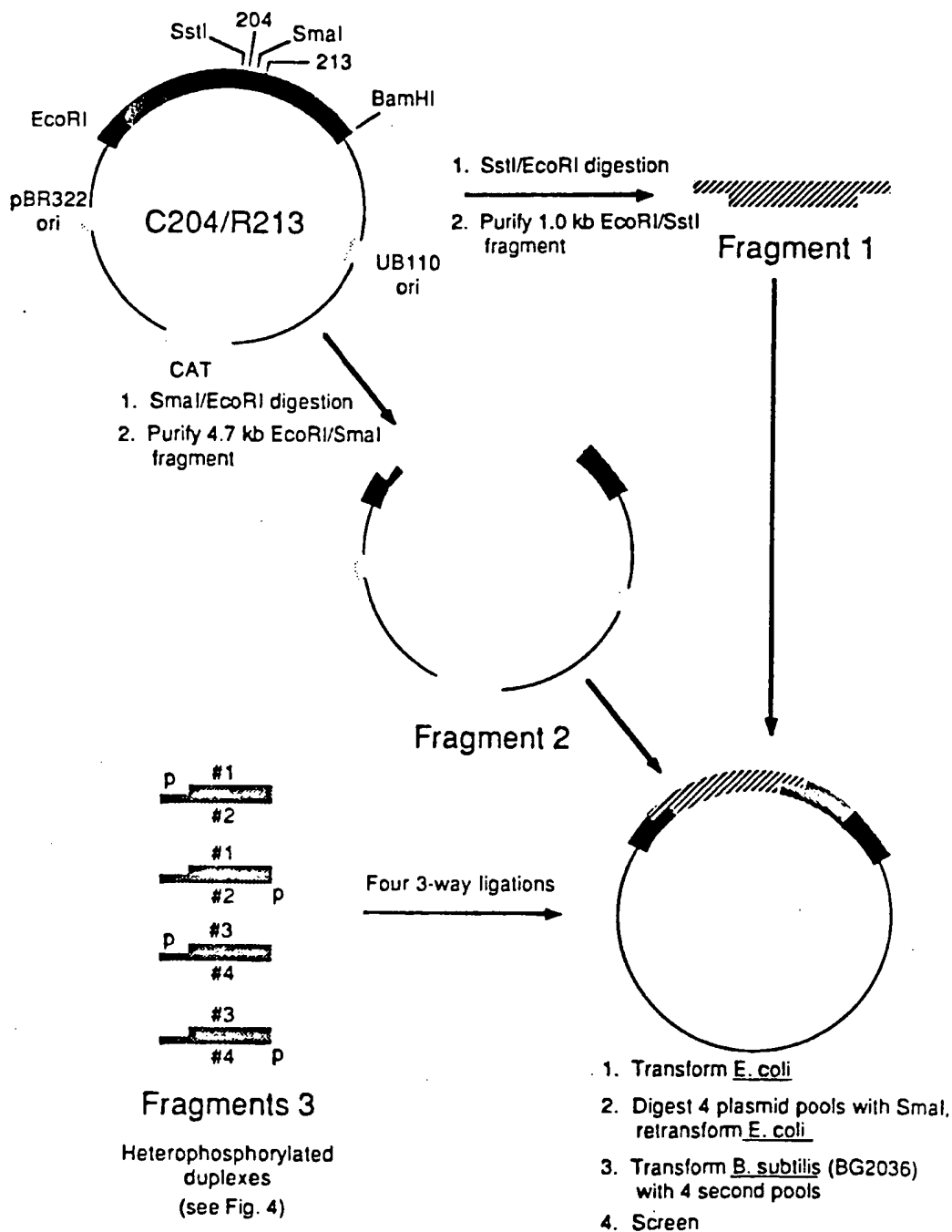


FIG.—36

Wild type A.A.:	195	200	204	210	213														
	Glu	Leu	Asp	Val	Met	Ala	Pro	Gly	Val	Ser	Ile	Glu	Ser	Thr	Leu	Pro	Gly	Asn	Lys
Wild type DNA:	5'-GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA AGC ACG CTT CCT GGA AAC AAA-3'	3'-CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT TCG TGC GAA CCT TTG TTT-5'																	
C204/R213 DNA:	5'-GAG CTC GAT GTC ATG GCA CCT GGC GTA TGT ATC CAA AGC ACG CTT CCC GGG AAC AGA-3'	3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT ACA TAG GTT TCG TGC GAA GGG CCC TTG TCT-5'																	
	SstI		SmaI																
C204/R213 cut with SstI and SmaI:	5'-GAG CT	3'-C	GGG AAC AGA-3'	CCC TTG TCT-5'															
C204/R213 cut and ligated with oligodeoxynucleotide pools:	5'-GAG CTC GAT CTC ATG GCA CCT GGC GTA	ATC CAG TCG ACG CTT CCT	GGG AAC AGA-3'	3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT	TAG GTC AGC TGC GAA GGA CCC TTG TCT-5'														
	SstI		SmaI																
	$\begin{array}{c} \text{W, R, R, or G} \leftarrow \text{NGG or} \frac{11}{12} \\ \text{Stop, Y, H, Q, N, K, D or E} \leftarrow \left[\frac{G}{C} \right]_{\text{TN}} \text{ or } \left[\frac{G}{C} \right]_{\text{AN}} \rightarrow \text{L, F, I, V or M} \end{array}$		$\begin{array}{c} \text{NCC} \rightarrow \text{S, P, T or A} \\ \left[\frac{G}{C} \right]_{\text{AN}} \rightarrow \text{L, F, I, V or M} \end{array}$																

FIG.—37